STREAMFLOW FORECASTING USING ARTIFICIAL NEURAL NETWORKS (CASE STUDY)

by

Hatim Mohammed Eisa

A thesis submitted in partial fulfillment of the requirements for the degree of M.Sc. in Water Resources Engineering

Department of Civil Engineering
University of Khartoum

Supervisor:

Dr. Babikir I. Barsi

November 2003

Abstract

Streamflow forecasting is an essential prerequisite to provide basic information on a wide range of flow-related activities and problems on natural and regulated rivers, such as, irrigation development, flood control, hydro-power generation, drinking water supply, and early warning systems.

Streamflow forecasts can be provided by the tool of hydrological modeling. Many Different types of models have been reported in the hydrologic literature. While conceptual or physically-based models are important in the understanding of hydrological process, there are many practical situations where the main concern is with making accurate predictions and forecasts at specific locations. In such situation it is preferred to implement a simple "black box" model to identify a direct mapping between the inputs and outputs without detailed consideration of the internal structure of the physical process.

This study presents the use of the recently developed modeling approach which is known as the Artificial Neural Networks (ANNs) to the problem of streamflow forecasting, and investigate its capability and applicability in simulating both rainfall-runoff as well as channel routing processes. The performance of the ANNs was then compared with other numerical simulators such as the Total Linear Model (TLM), the Modified Linear Perturbation Model (MLPM), and the Autoregressive (AR) Model.

All the models were applied to the medium sized (166,875 km^2, 8°-15°N, 34°-39°E) Atbara River sub-basin which is part a of the Nile River basin. Six years of real data of lumped daily rainfall and streamflow were used during the study.

The results indicate that the nonlinear ANN model approach provides superior performance over all models across the full range of flow levels. Consequently, these results suggest that the ANN approach may provide a superior alternative to the Total Linear Model (TLM) and the Modified Linear Perturbation Model (MLPM) approaches for developing input-output simulation and forecasting models in situations that do not require modeling of the internal structure of the watershed.

TABLE OF CONTENTS

Chapter 3: The Fundamentals of the Models Used During the Study

Chapter 4: Model Data Development

Chapter 5: Calculation, Results and Discussion

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

The author is greatly indebted to his supervisor Dr. Barsi for his encouragement and invaluable advice in conducting the research and writing the thesis.

Acknowledgements are due to Dr. Gamal, Dr. Adil and Eng. Asim for their help and fruitful discussions.

The author is grateful to the members of the staff of Civil Engineering Department and his colleges for the friendly atmosphere during the course of study.

Special thanks are due to the A. Sighairoon, M. Al-Mostafa, Igbal and Taghreed for their sincere help in supplying the required data and fruitful discussions.

C h a p t e r 1

INTRODUCTION

1.1 General:-

Water plays a critical role in maintaining the well being of the society and the natural systems. However, water problems and stresses are complex and vary with causes, type, time, and location. Among them, flooding is a major natural disaster that has been a significant threat to economic development and human life in many countries. To address the problem of floods it is important to clarify the characteristics of flooding. For this purpose floods may be viewed from a natural (hydrological and ecological) perspective and a human perspective.

What is a flood? Viewed from a hydrological perspective, a flood may be described as an abnormally high stage in a river channel and its floodplain. The type of flood situation varies with location along the river course. Three main factors govern the dynamics and development of a natural flood event in a river valley: (1) intensive storm rain fall and /or rabid snow melt; (2) watershed characteristics and prevailing conditions influences the runoff regime (the type of basin, vegetal cover and its modification, the antecedent moisture content); and (3) river channel and flood plain conditions (Davar et al, 2001).

From the human perspective of those who have occupied and developed the flood plain, a major flood event is one that at high river levels results in extensive damages (economic), general disruption of human activities (social effects), and intense trauma (psychological). The attractiveness of a river floodplain during non-flood seasons leads to its occupancy and development by humans.

Over time and with much experience, humanity has developed and practiced a variety of defensive measures in response to the periodical threat of floods. Often, these approaches have been used in combination in response to the needs of a specific flood

situation. Techniques for flood damage reduction encompass structural and non-structural measures that reduce or eliminate flood hazard and the potential for future damages. These measures are listed in Table (1.1) and shown in Figure (1.1).

So the current study can be classified as one of the non-structural measures for flood damage reduction.

1.2 Background and Motivation:-

The Nile is one of the world longest rivers, flowing south to north 6,850 kilometers, over 35 degrees of latitude (3°S-32°N), (AmaraseKera et al, 1996), Table (1.2) (Karyabwite, 2000) and Table (1.3) shows a brief comparison of some °hydrological characteristics of the world's major river systems.

The catchment of the Nile basin covers approximately 10 % of the African continent, with an area of $3*10^6$ Km^2, (Karyabwite, 2000), annual precipitation of about 1900 km^3, a ratio of precipitation to catchment area of about 0.64 m (Amarasekera et al, 1996), and spreads over 10 countries, Table (1.4) (Karyabwite, 2000) and Figure (1.2) (Georgakakos, 1996) shows the Nile basin countries.

The shape of the Nile today is a very recent development. The present day river is complex and is the result of the interconnection of several independent basins by rivers which developed during the last wet period which affected Africa after the retreat of the ice of the last glacial age, some 10,000 years ago. The basin of the Nile is characterized by the existence of two mountainous plateaus rising some thousands of meters above (MSL).

Table (1.1): Flood damage reduction measures (Davar et al, 2001)

Structural Measures	
lood Diversion	lood Water Diversion Channels
	lood Water/ice debris storage area

Channel structures	Storage dams and reservoirs
	Diversion dams and barrages
	Ice or debris control structures
	Flow retardation basins
Flood barriers	Ring Dykes and polders
	Levees and flood walls
Structural flood proofing	Permanent flood proofing (requiring no action prior to flood) elevation of structure on fill or columns berms, cut-off walls flood proofing of utilities flow (e.g., check valves on sewers) water-resistance materials and construction Contingency flood proofing (requiring action just prior to flood) watertight doors flood shields
Non-Structural measures	
Emergency measures	Flood fighting/emergency/flood proofing (e.g., placing sand bags)
	Evacuation
	Disaster relief services (e.g., water and food distribution)
Flood advisories	Flood alerts (based on long-term and seasonal information)
	Flood warning (based on upstream observed/predicted conditions)
	Flood forecasts (based on hydrologic analysis/modeling)
Flood plain management	Public education and awareness programs
	Flood plain delineation Flood mapping Flood plain area demarcation past flood level indicators/signs
	Financial measures Financial incentive /disincentives Flood insurance
	Direct change relocation use conversion
	Regulation of land use activities National/state/provincial legislation Regional or municipal zoning by-laws Building covenants and restrictions subdivision control restrictions on issuance of development permits watershed management to reduce flood volumes and runoff rates Other land use controls and planning

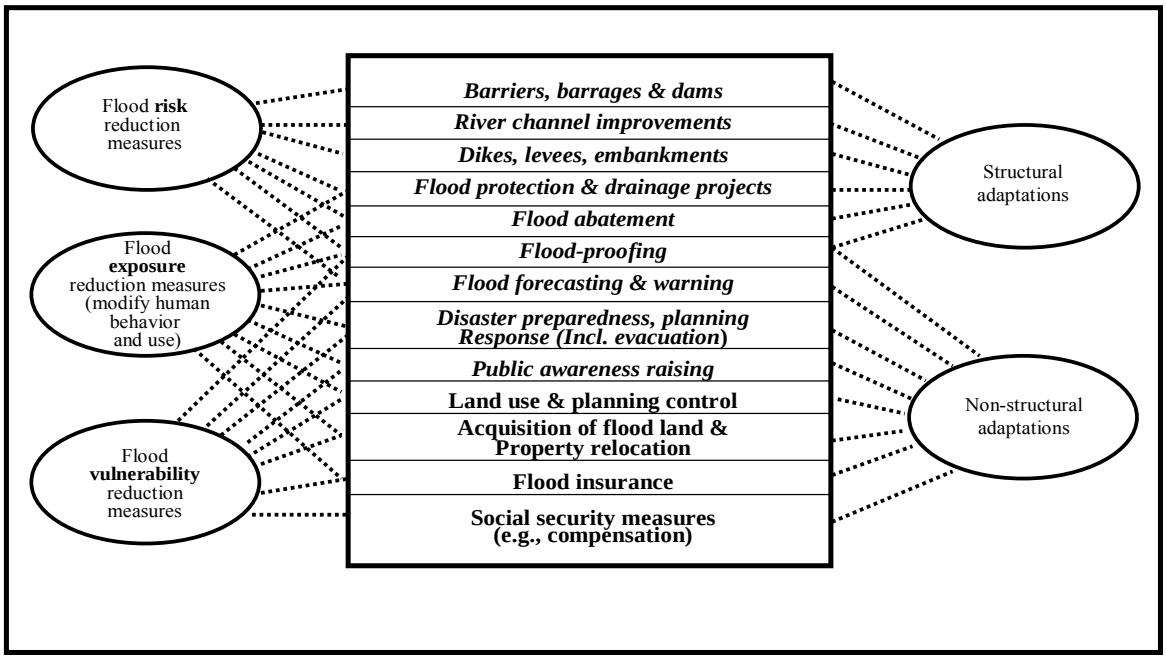


Figure (1.1): The range of options to mitigate inland flooding (Bourget, 2001)**.**

Table (1.2): World's major river systems (Karyabwite, 2000).

River	Length (km)	Drainage area (10^3 km3)	Annual discharge (10^9 m^3)	Discharge/unit area (10^3 m^3/km^2)
Nile	6850	3110	84	28
Amazon	6700	7050	5518	728
Congo	4700	3820	1248	326
Mekong	4200	795	470	590
Niger	4100	2274	177	78
Mississippi	970	3270	562	170
Danube	2900	816	206	252
Rhine	1320	224	70	312
Zambezi	2700	1200	223	185

Table (1.3): Some hydrological characteristics of some world's major river systems (Amarasekera et al, 1996)

River	Length (m)	Drainage area (x $10^6 km^2$)	Annul runoff (km^3)	Latitude extent of drainage basin	Annual precipitation (km^3)	Runoff/ catchment (m)	Precipitation/c atchment (m)
Amazon	6577	6.16	6300	4°N-18°S	12000	1.02	1.95
Congo	4375	3.82	1250	7°N-12°S	5600	.33	1.47
Paraná	3740	2.6	470	16°S-34°S	3300	0.18	1.27
Nile	6648	2.96	83	3°S-32°N	1900	.03	0.64

Table (1.4): Nile River basin countries (Karyabwite, 2000)

Country	Total area of the country (km2)	Area of the country within the basin (km2)	As % of total area of basin (%)	As % of total area of country (%)	Average annual rainfall in the basin area (mm)		
					Min.	Max.	Mean
Burundi	27,834	13,260	0.40	47.60	895	1,570	1,110
Rwanda	26,340	19,876	0.60	75.50	840	1,935	1,105
Tanzania	945,090	84,200	2.70	8.90	625	1,630	1,015
Kenya	580,370	46,229	1.50	8.00	505	1,790	1,260
Zaire	2,344,860	22,143	0.70	0.90	875	1,915	1,245
Uganda	235,880	231,366	7.40	98.10	395	2,060	1,140
Ethiopia	1,100,010	365,117	11.70	33.20	205	2,010	1,125
Sudan	2,505,810	1,978,506	63.60	79.00	0	1,610	500
Egypt	1,001,450	326,751	10.50	32.60	0	120	15
For Nile basin	8,919,072	3,112,369	100	34.90	93.22	1528.30	602.96

The Equatorial or Lake Plateau in the southern part of the Nile basin, situated between the two branches of the Great Rift, is at a level of 1,000 to 2,000 meters and has peaks of

5,100 and 4,300 meters above (MSL). This plateau contains Lakes Victoria, George, Edward and Albert, which slope gently toward the north at an average rate of one meter for every 20 to 50 km of stretch. In contrast the rivers which connect these lakes fall at an average rate of one meter every kilometer or less of length.

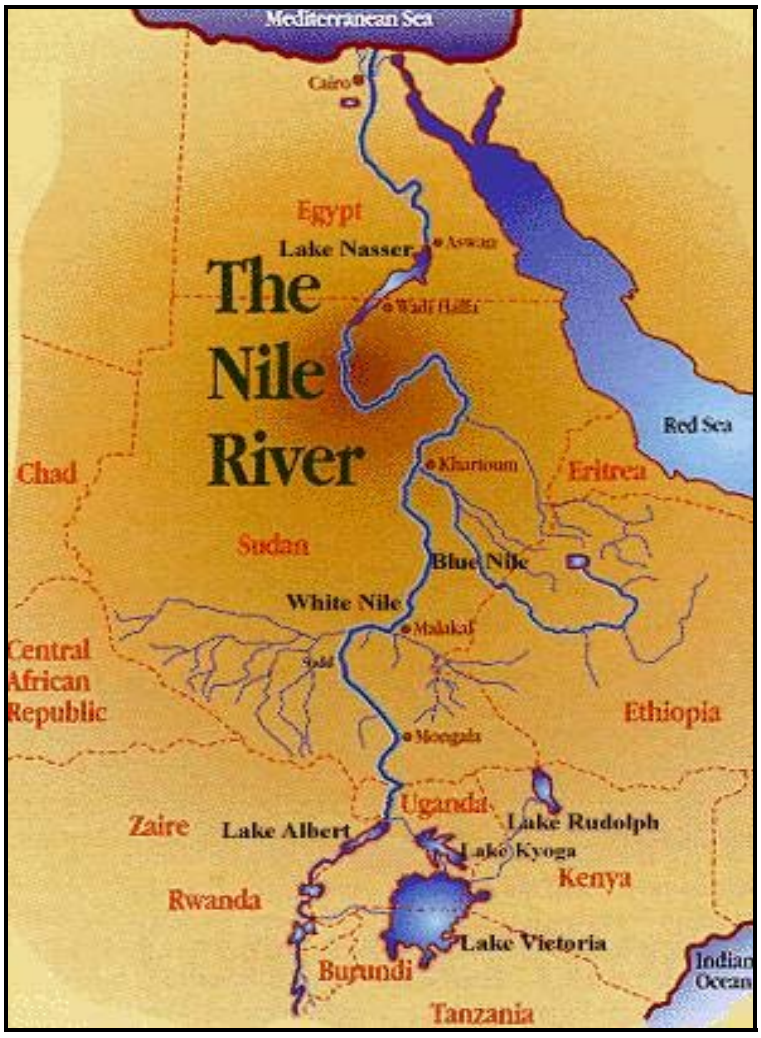


Figure (1.2): Nile River countries map (Georgakakos, 1996)

The Ethiopian or Abyssinian Plateau, which forms the eastern part of the basin, has peaks rising to 3,500 meters. North of the Lake plateau the basin descends gradually to the Sudan plains where the Nile runs at altitudes lower than 500 m in its northerly direction. About 200 km south of the Egyptian border the river cuts its channel in a narrow trough

bounded from each side by the contour line of 200 m ground surface level. Almost 200 km before discharging into the Mediterranean sea, the river bifurcates and its two branches encompass the Nile Delta (Karyabwite, 2001)

In order to have a fair insight into the flood wave formation of the Nile, especially within Sudan, more attention must be given to the Blue Nile and Atbara rivers which together contribute more than 90% of the peak flows in the central arid regions. In fact the White Nile basically contributes about 80-90% of the low flows and around 30% of the annual flow volume (Andah et al, 1991), (Figure (1.3)). The Blue Nile which emerges from the Lake Tana in the Ethiopian Plateau produces very strong torrential flows during the rainy season and exceptionally low flows in the dry periods. The ratio of the maximum flows to the minimum can reach a very high value of 20: 1. The runoff generation is governed by the humid climate conditions in the Ethiopian Highlands in conjunction with its sharp topographic characteristics. From Lake Tana to the Roseires Dam, about 1000 km downstream, the channel falls steeply with a slope of about $1.3*10^{-3}$. analysis of peak flows at Khartoum indicates that floods in and around Khartoum are generated by the Blue Nile, while at the same time causing high backwater flows into the White Nile. In the period July-August, the ratio of the flows of the Blue Nile to those of the White Nile is in the order of 97:1. It must however be noted that the flow from Roseires to Khartoum is gentler due to channel slopes of $1.2\ 10^{-4}$ with some minor routing effects of the Roseires reservoir (Andah et al, 1991).

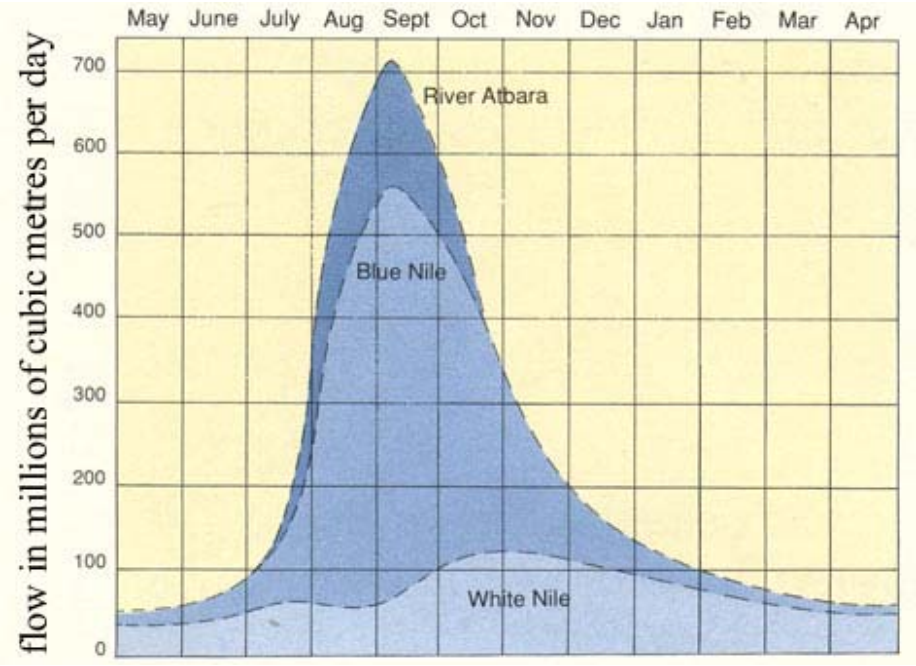


Figure (1.3): The cumulative hydrograph of the White Nile, Blue Nile, and the Atbara River (Barron, 1998).

The most important subcatchment of the Main Nile is the Atbara river basin which has hydrological and morphological characteristics similar to the Blue Nile. The river Atbara also emerges from the Northern highlands of Ethiopia with an initial steep slope of $5*10^{-3}$, carrying torrential high flows and large amounts of sediments during the rainy season. The Atbara is highly seasonal and most often dries up to 5 months. In exceptionally wet years, the flows can reach about $370 *10^{6}$ m^3/day, measured through the Khashem El-Girba dam. It must be noted that the effective base width of the annual hydrograph comprises June-December with the flood season covering only August and September. After Atbara there is no other significant inflow to the Main Nile.

As a consequence of the historical (1988) extensive flooding event, there were serious widespread damages to agricultural farms and pumping installations along the Nile valley, especially the northern areas of Atbara and Dongola. Most of the islands situated in the Main Nile were inundated or cut away from the rest of the country with grave consequences to life and property. The most serious socio-economic effects of the flood

events were the complete loss of the year's farming season. The Northern provincial towns of Atbara, Dongla, New Halfa and many others were heavily hit by both the Nile floods and the causative effects of high intensive rainfalls on urban inundation and flash floods through temporary natural channels.(Andah et al, 1991).

3-Objectives and Scope of the Study:-

The objective of the current study is to develop a hydrological model allowing short-term simulation and application for real time streamflow forecasting in order to provide an early warning system with a reasonable lead time downstream. This will reduce the flood damages, and enable efficient watershed management.

Streamflow forecasts were provided by:-

1- Simulating the rainfall-runoff process over the study area to forecast the inflows into two main tributaries of Atbara River namely: Setit, and river Atbara.

2- Simulating the channel routing process to forecast the inflow to Khasm El-Girba dam, this is generated as a result of the combined effects of the hydrographs of Setit and river Atbara.

3- Simulating the channel routing process of the release from Khasm El-Girba dam and the streamflows downstream at the mouth of Atbara River.

In order to provide the streamflow forecasts, four different simple black box models were investigated. The models are: the Artificial Neural Networks (ANNs); which is a recently developed modeling approach, the Total Linear Model (TLM); which was developed by Liang and Nash in 1988 (Abdo et al, 1992), the Modified Linear Perturbation Model (MLPM) whish is based on the Linear Perturbation Model (LPM)developed by Nash and Barsi in 1983 (Abdo et al, 1992), and the Autoregressive Model (AR); which its basic theory is due to Yule in 1927and Wolker in 1931 (Kottegoda, 1980).

1.4 Thesis Organization:

The text contained in the chapters from 1 to 7, and they were organized as follows:

Chapter 2 gives briefly general information about the science of hydrology as well as the art of hydrological models and their classification. Also, a short review about the models applied to the region under study was reported.

Chapter 3 gives, in detail, the fundamentals of the models used during the study namely; Artificial Neural Networks (ANNs), Linear Perturbation Model (LPM), the Modified LPM, the Total Linear Model, and the Autoregressive model (AR), with more stressing on the recently developed modeling approach whish is the ANNs.

Chapter 4 describes the data used during the study. It contains a general description of the region considered during the study and how the river network and watershed discretization were identified. Quality of rainfall data is questioned and tested for homogeneity. Missing stage amounts were determined and different models for filling the in this amounts were suggested and compared. A procedure for filling in missing stage amounts is outlined and applied. A simple water balance model for determining the inflows to Khashm El-Girba reservoir was mathematically detailed and applied.

Chapter 5 concerns with the application and identification of the models, descried in chapter 3, to the study area. The model components were briefly described. The building process of the ANNs model considering the architecture of the network was given in detail. The results were also presented and discussed.

Chapter 6 presents the summary and conclusions.

All the appendices are related to chapter 4.The total number of appendices is 6.

C h a p t e r 2

LITERATURE REVIEW

2.1 General: -

Modern society faces a series of challenges, one of the most important concerns the uses and abuses of water. Few natural resources are as universally and as constantly in demand, and few are employed in so many ways. Water is needed for drinking, for power generation and for cooling purposes; it is required for the carriage of wastes and for irrigation; it is the basic for many leisure activities. Too much water causes destruction and leads to diseases; too little results in shortage and famine. Intense erosion can lay waste fertile regions; serious pollution destroys the life of rivers and can bring about epidemics. To combat these extremes large sums of money are spent in many countries on the construction of dams and flood alleviation schemes, on the development of water distribution systems and on the provision of purification works. Basic to many water resources projects is hydrology, a science with it is foundations in ancient Egypt and Mesopotamia and one that can be approached in a number of different ways. Engineers, chemists, physicists, geologists, meteorologists, geographers and biologists have all made significant contributions to the science of hydrology. Application of the principle of hydrology creates a potential for improving the lives of those in the world without an adequate water supply. In countries where water resources are almost fully developed and there are the conflicting demand of water supply, effluent disposal, recreation, amenity and environmental, hydrology provides a framework for planning and development.

Hydrology can be defined as the science that treats the waters of Earth, their occurrence, circulation and distribution their chemical and physical properties, and their reaction with their environment, including their relation to living things. The hydrological cycle see Figure (2.1), which is the most fundamental principle of hydrology, can be described as follows: water evaporates from the oceans and the land surface, is carried over the earth

in the atmospheric circulation as water vapor, precipitates again as rain or snow, is intercepted by trees and vegetation, provides runoff on the land surface, infiltrates into soils, recharges groundwater, discharges into streams, and ultimately, flows out into the oceans from which it will eventually evaporates once again. This immense water engine fueled by solar energy, driven by gravity, proceeds endlessly in the presence or absence of human activity (Chow et al, 1988).

2.2 History of Hydrology: -

The beginnings of hydrology are lost at the start of history: the science is at least as old as the ancient riverine civilizations of Egypt, the Indus and Mesopotamia. Nile flood records date from before 3000B.C., while the early irrigation systems reached peaks of development about 2000B.C. (Rodda et al, 1976) Frontinus, Water Commissioner of Roman in 97A.D.-based estimates of flow on cross-sectional area alone without regard to velocity.

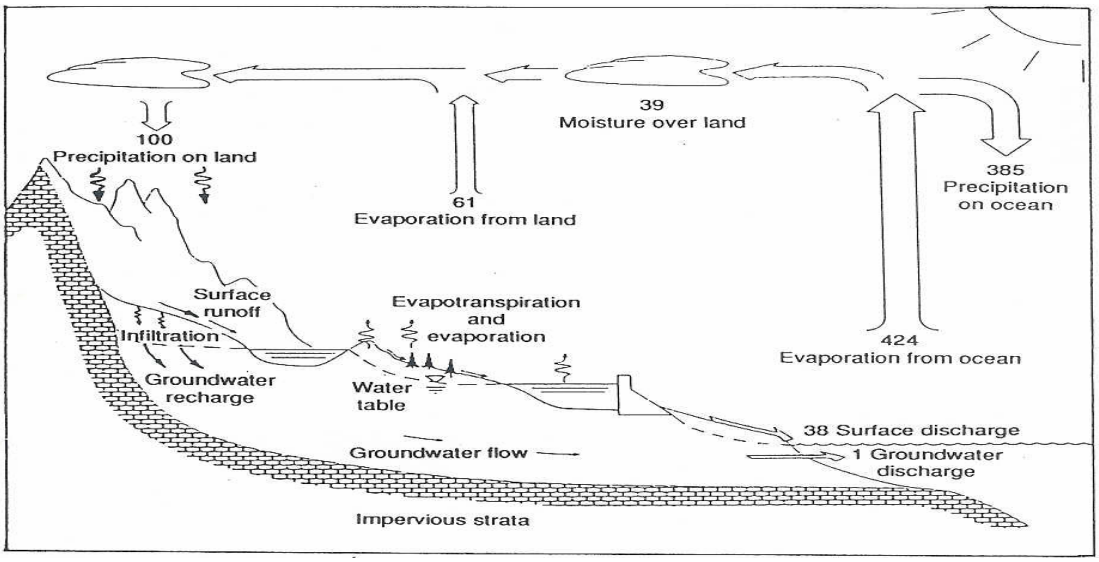


Figure (2.1): The hydrological cycle with annual volumes of flow in units relative to the annual precipitation on the land surface (119,000 km^3/year) (Chow et al, 1988).

In the United State, organized measurement of precipitation started under the Surgeon-General of the Army in 1819, was transferred to the Signal Corps in 1870, and finally to

a newly organized U.S. Weather Bureau in 1891. It is not surprising; therefore, that little quantitative work in Hydrology was done before the early years of the twentieth century, when men such as Horton, Mead and Sherman began to explore the field. The great expansion of activity in flood control, irrigation, soil conservation, and related fields, which began about 1930, gave the first real impetus to organized research in hydrology, as the need for more precise design data became evident. Most of our present day concepts of hydrology date since 1930. Hydrology is therefore, a young science with many important problems only imperfectly understood and much research still a head.

Ven T Chow (Varshney, 1979) divides the history of hydrology into eight distinct periods as follows:

1- Period of speculation (up to 1400A.D.)
2- Period of observation (1400-1600A.D.)
3- Period of measurements (1600-1700A.D.)
4- Period of experimentation (1700-1800A.D.)
5- Period of modernization (1800-1900A.D.)
6- Period of empiricism (1900-1930A.D.)
7- Period of rationalization (1930-1950A.D.)
8- Period of theorization (1950-

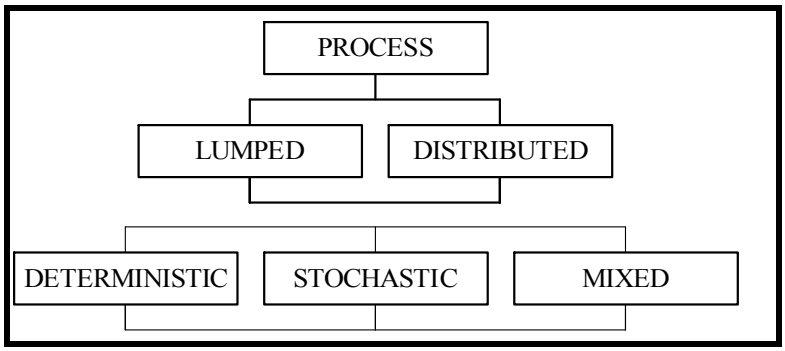
Inspite of considerable headway made by the science of Hydrology, it is not very much recognized as a science in many quarters and at best considered as a science of coefficients and empirical formulae. Even where some concepts of hydrology are known, some people still feel reluctant to discard reliance on empirical formulae (Varshney, 1979).

2.3 The Art of Hydrological Modeling: -

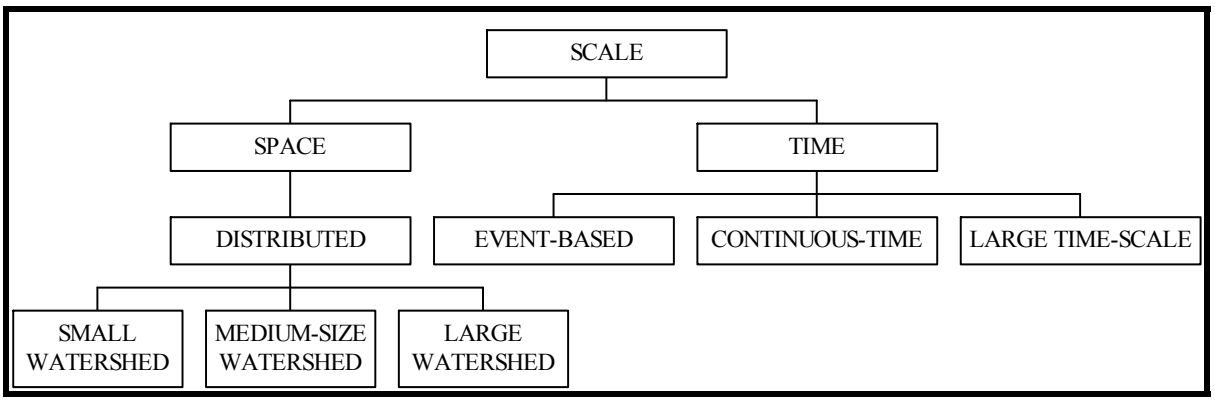
A model is a simplified representation of a system or process in which only the basic concept are considered. It is used to identify relationships and interrelationships in terms of cause and effect. In essence a model describes the conservation of one phenomenon

into another, an 'input' into an 'output', in a situation where continuity in time and in quantity is important (Rodda, et al 1976).

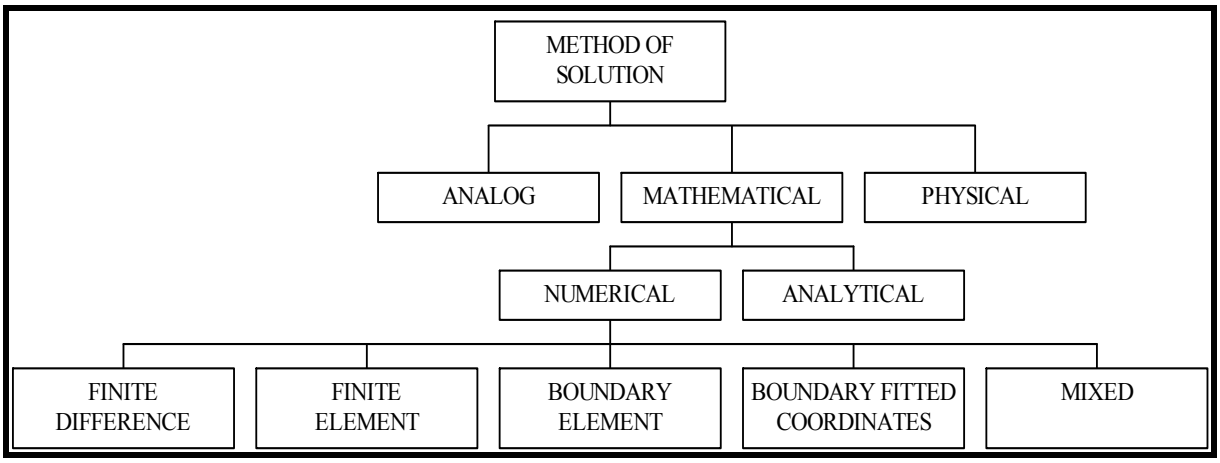
The models are different in type and where developed for different purposes. Nevertheless, many of the models share structural similarities, because their underlying assumptions are the same, and some of the models are distinctly different. The watershed models can be classified, as shown in Figure (2.2), according to different criteria that encompass (1) process description, (2) scale, and (3) technique of solution (Singh, 1995).



(2.2a) Classification of models based on process description



(2.2b) Classification of models based on space and time scales.



(2.2c) Classification of models based on solution technique.

Figure (2.2): Classification of watershed models (Singh, 1995).

2.3.1 Process-Based Classification: -

A model as shown in Figure (2.3) has five components, including (1) system (watershed) geometry, (2) input, (3) governing laws, (4) initial and boundary conditions, and (5) output. Depending upon the type of the model, these components are variously combined. The processes include all of the hydrologic processes that contribute to the system (watershed) output. Based on the description of those processes, conjunction with the system characteristics, the models can be described as lumped or distributed, deterministic or stochastic or mixed as shown in Figure (2.2a).

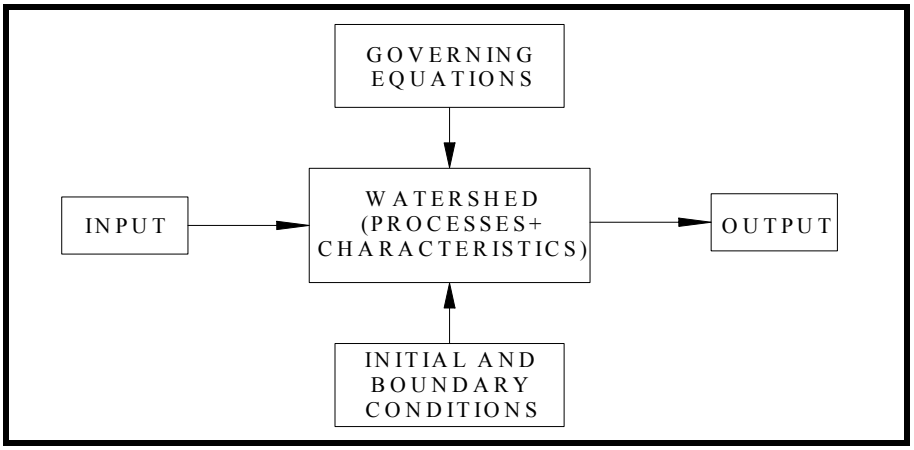


Figure (2.3): Model components.

A lumped model, as shown in Table (2.1), is, in general, expressed by ordinary differential equations taking in to account of spatial variability of processes, input, boundary conditions and system geometric characteristics. In most lumped models, some processes are described by differential equations based on simplified hydraulic laws, and other processes are expressed by empirical algebraic equations.

Table (2.1): Lumped and distributed models (Singh, 1995).

Input	System Characteristics	Component Processes	Governing Equations	Output	Model Type
Lumped	Lumped	Lumped	ODE	Lumped	Lumped
Lumped	Lumped	Distributed	PDE	Distributed	Distributed
Distributed	Distributed	Distributed	PDE	Distributed	Distributed
Distributed	Lumped	Distributed	PDE	Distributed	Distributed

Distributed models, as shown in Table (2.1), take an explicit account of spatial variability of processes, input, boundary conditions, and/or system characteristics. Of course, in practice, a lake of data—field experimental (laboratory)—prevents such a general formulation of distributed models. In a majority of cases the system characteristics are lumped, many of the processes are lumped, the input is lumped, and even some of the boundary conditions are lumped, but some of the processes that are directly linked to the output are distributed, as for example rainfall-runoff process. These models are not fully distributed; rather they are quasi-distributed at best.

The description of a process can be deterministic, stochastic or mixed. Depending upon the way the processes are described, the models can be also classified as deterministic, stochastic, or mixed. For a deterministic model, the output is uniquely determined by the input. On the other hand, a stochastic model produces an output that contains time-dependant random variable. If all of the components of a model are deterministic, the watershed model is deterministic. If all of the components of the model are stochastic, the

model is fully stochastic, and if only some components are stochastic then the model is quasi-stochastic. When the model components are described by a mix of deterministic and stochastic components, the model is a stochastic-deterministic or hybrid model (Singh, 1995). Watershed models can also be classified into conceptual and empirical according to whether they are, or aren't, capable of physical interpretation. Conceptual models are developed by purely rational consideration involving the interplay of inductive and deductive reasoning. The conceptual model's parameters are either physically measurable or can be interpreted in physical terms. However, the empirical approach may prove quite satisfactory and valuable for decision-making.

2.3.2 Time-Scale Based Classification: -

The watershed models can be classified based on the time scale of the models. The time scale can be defined as a combination of two-intervals. One of the time intervals is used for input and internal computations. The second is the time interval used for the output and calibration of the model. Based on this description, the models can be distinguished as (a) continuous-time or event based, (b) daily, (c) monthly, and (d) yearly models. This classification is dictated by the interval of computation. If the model components are available at shorter time intervals, such as hourly interval or shorter, the model will be a short interval model. Of course the choice of a time interval is often a function of model's intended use (Singh, 1995).

2.3.3 Space-Scale Based Classification:-

The spatial scale can be used as a criterion to classify models into small watershed, medium-size watershed, and large watershed models. Of course, the definition of small watershed is somewhat vague. Usually, the watershed with area of 100 km² or less can be called small, those with area of 100-to 1000 km² medium, and those with area of larger than 1000 km² large. Such a classification is arbitrary, and is experimental rather than conceptual and is governed by data availability rather than physical meaning. The essential important is the homogeneity and averaging of hydrological processes. For consideration of runoff generation on these watersheds, two phases can be considered:

land phase and channel phase. Each phase has it's own storage characteristics. Large watersheds have well developed channel networks and channel phase, and thus, channel storage is dominant. Such watersheds are less sensitive to short duration, high intensity rainfalls. On the other hand, small watersheds have dominant land phase and overland flow, have relatively less conspicuous channel phase, and are highly sensitive to high-intensity, short-duration rainfalls.

If all three types of watersheds are spatially uniform, then they will behave hydrologically similarly. Under this condition, classification of watersheds serves no useful purpose and becomes moot (Singh, 1995).

2.3.4 Solution Technique Based Classification: -

Based on solution technique, the models can be classified into physical, analog and mathematical. Physical models have often been used to investigate the behavior of hydraulic structures and hydrological system involving rivers and estuaries.

An analog model is a mechanical or electrical device whose operation is governed by the same fundamental equations as those of the system being modeled; thus there exists quantitative relationships between variables in the model and in the prototype. Mathematical models simulate the behavior of the system through a set of computer programs and mathematical relationships; it can also be sub-classified into analytical and numerical as shown in Figure (2.2c).

2.3.5 Land-Use Based Classification: -

Based on land-use, the watersheds may be classified as (1) agricultural, (2) urban, (3) forest and range land, (4) desert, (5) mountainous, (6) coastal, (7) wetlands, and (8) mixed. In any cases, large or even medium-size watersheds have mixed land use. These watersheds behave hydrologically differently, indeed so differently that they have given rise to different branches of hydrology. For example, treatment of agricultural watersheds constitutes what is called agricultural hydrology. Similarly there is urban hydrology for

urban watersheds, coastal hydrology for coastal watersheds, forest hydrology for forest watersheds, desert hydrology for desert watersheds, wetland hydrology for marshes and wetlands, and mountain hydrology for mountainous watersheds. The soils, geology, and vegetation are different for these watersheds. As a result, the hydrologic processes are different in their evolution in these different watersheds, and hence their models are clearly different (Singh, 1995).

2.3.6 Model-Use Based Classification: -

Hydrological models can be applied as decision-making tools, particularly in to two aspects of prognostication. The first aspect, usually called forecasting, which involves the estimation of the value of some hydrological variable at a prescribed instant of time or a forecast of the time when a particular value of a variable will occur. The second aspect usually known as prediction which is concerned not with the value of a single event or its time of occurrence, but with the frequency of occurrence of some prescribed critical condition, for example, how often a particular flood level will be exceeded.

Frequently, watershed models are classified on the basis of their intended use: (1) planning models, (2) management models, and (3) prediction models. A comprehensive watershed model can be employed to accomplish a considerable array of analytical tasks for planning and management of water resources. It offers the opportunity for systematic analysis of a management policy. On the other hand, it becomes an increasingly valuable predictor of river-basin response to management strategies. For planning purposes, the model defines geographical and temporal trends that may not be accurate at specific site in the river basin. Thus, the primary use of a planning model is for analyses of river systems management strategies. Probably the most powerful feature of planning and management model is the ability it offers planners and managers to replace past management indices with improved economic performance of potential decisions (Singh, 1995).

2.4 Perspective on Model Building: -

Construction of a computer model of watershed hydrology may be driven by the need to solve a particular problem or the scientific pursuit of model building, which may or may not be used for problem solving. Frequently, the type of a model to be built is dictated by the availability of data. Different models have different data needs. In general, distributed models require much more data than do lumped models. In most cases, needed data either do not exist or are not available in full. That is why regionalization and synthetic techniques are useful. Even if the needed data are available, problems remain with regard to incompleteness, inaccuracy, and in-homogeneity of data. Then, of course, storage, handling, retrieval, analysis, and manipulation of data have to be dealt with. If the volume of data required is large, data processing can be quite a sophisticated undertaken.

A watershed hydrology model is an assemblage of component models corresponding to different components of hydrologic cycle. The time-scale of model output (e.g. streamflow) greatly influences the type of the model or details to be included in the model. For example, a monthly watershed model is quite different in its architecture and construction from, say, an hourly model. Thus, the elements of simulation differ with the model type. It remains an unresolved question as to hydrologic laws operating at different times-scales for different components of the hydrologic cycle. Solution of this question will greatly facilitate model construction and more clearly define the data needs.

After a model is developed, it needs calibration, verification, and an assessment of its reliability. Model calibration involves optimum parameter estimation subject to a specific error criterion. Usually the available data is split into two halves, one for calibration and the other for verification. The calibration set should encompass the full range of data available that may be expected in verification. However, it is not clear how model's parameter change with variation in the objective function. Furthermore, an error analysis of calibration results is to be performed. Through this analysis, it can be ascertained if the errors are due errors in data, parameters, or model construction. Model verification and reliability of model results are essential to show if the model is any good (Singh, 1995).

2.4.1 **Model Calibration: -**

However, model results are only as reliable as the model assumptions, inputs, and parameter estimates. Therefore, two problems had been face. The first is to select suitable model to represent the study site (watershed). The second is to then select values for the model parameters so that the model closely simulates the behavior of the study site. The process by which the parameters are selected is called model "calibration".

A typical parameter estimation procedure consists of four major elements: (1) objective function, (2) optimization algorithm, (3) termination criteria, and (4) calibration data (Singh, 1995).

2.4.1.1 ***Objective Function*: -**

An objective function is an equation that is used to compute a numerical measure of the difference between the model-simulated output and the observed output. The purpose of model calibration is therefore: "to *find those values of the model parameters that optimize (minimize or maximize, as appropriate) the numerical value of the objective function*".

In this study the Simple Least Squares (SLS) function is used as an objective function:

$$F(\beta) = \sum_{t=1}^{n} \left[q_t^{obs} - q_t(\beta) \right]^2 \quad \text{.....(2.1)}$$

Where:

$q_t^{obs} \equiv$ Observed (measured) value at time t.

$q_t(\beta) \equiv$ Model simulated value at time t.

$\beta \equiv$ Vector of model parameter.

$n \equiv$ The number of data points to be matched.

And assume that the objective function must be minimized.

2.4.1.2 *Optimization Algorithm*: -

The surface described by the objective function in the parameter space is called a "response surface". An optimization algorithm is a logical procedure that is used to search the response surface, constrained to the allowable ranges on the parameters, for the parameter values that optimize the numerical value of the objective function. The gradient search optimization strategy is used in this study, which is categorized as one of the local search methods. Most gradient methods are based on the following equation:

$$\beta_{i-1} = \beta_i - \rho.\Delta f(\beta_i) \quad \text{.....(2.2)}$$

Where:

$\beta_i \equiv$ The present (initial) point parameter vector.

$\Delta f(\beta_i) \equiv$ The function gradient matrix at the present point.

$\rho \equiv$ A step size parameter.

$\beta_{i-1} \equiv$ The new point parameter vector.

If the exact matrix of second partial derivatives with respect to the parameters, the optimization method would be the well known "Newton method" exactly the quasi-Newton (refer to Gill et al, 1991 for more details) which was used in this study. During the study an optimization algorithm supports to the Microsoft spreadsheet was used, which is called "Solver", see Figure (2.4).

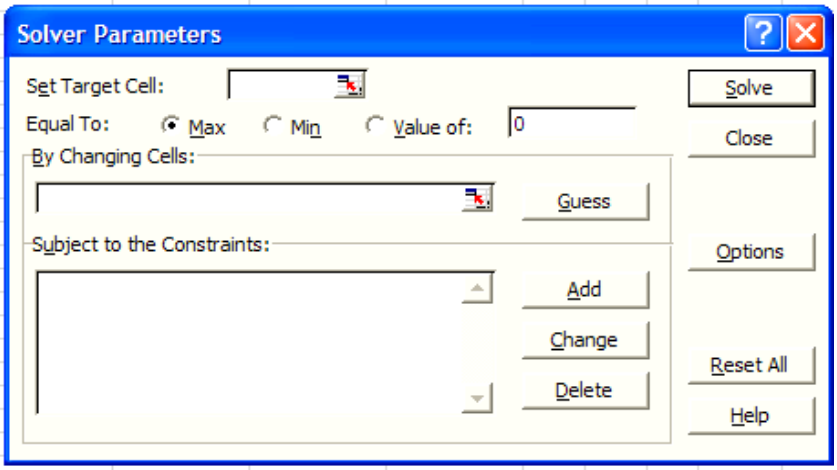


Figure (2.4): Microsoft spreadsheet's optimization algorithm.

2.4.1.3 *Termination (stopping) Criteria*: -

The optimization strategy used is iterative procedure, which search for the optimal parameter values by means of incremental improvement steps. Therefore, a criterion is needed to determine when to stop the search. In principle, the solution exists at that point in the parameter space where the slope of the function response is zero and the function value is a minimum. In practice, it is virtually impossible to know when this point has been reached; hence the following criterion is used:

$$(f_{i-1} - f_i) / f_i \leq \varepsilon_f \quad \text{.....(2.3)}$$

Where:

f_{i-1} & $f_i \equiv$ The function values at the $i-1^{th}$ and i^{th} steps, respectively.

$\varepsilon_f \equiv$ The function convergence criterion (e.g. 10^{-3}).

2.4.1.4 *Statistical Interpretation*: -

The model performance was evaluated on both qualitatively by the visual inspection of the simulated and the observed hydrographs, and scatter-plots, and quantitatively using the following statistical measures: Sum of Square Error (SSE), Standard Error of Estimate (SEE.), F test of significance with 95% confidence level (F-value), coefficient

of determination (R^2), Root Mean Squared Error (RMSE), Error Bias which measures the percent error in volume(bias)(Hsu et at,1993), the percent error in matching the maximum (peak) flow (%MF)(Hsu et al 1993), and Model Efficiency (E_{ff}) which is developed by Nash and Sutcliffe in 1970 (Dibike et al,1999), and can be mathematically expressed as

$$SSE = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \quad \text{.....(2.4)}$$

Where:

$(Y_i - \hat{Y}_i) \equiv$ The deviation of the i^{th} observation from its predicted value (the i^{th} residual).

$$SEE = \sqrt{(SSE/df)} \quad \text{.....(2.5)}$$

Where:

$df \equiv$ Degrees of freedom of the residual sum of squares.

$$E_{ff} = \frac{(F_o - F_1)}{F_o} \quad \text{.....(2.6)}$$

Where:

$F_o \equiv$ The initial variance of the observed amounts about the mean of the series.

$$F_o = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \quad \text{.....(2.7)}$$

$\bar{Y} \equiv$ The mean of the observed values.

$(Y_i - \bar{Y}) \equiv$ The deviation of the i^{th} observation the series mean.

$F_1 \equiv$ The sum of squared residuals or (SSE).

$$F = \frac{MS}{S^2} \quad \text{.....(2.8)}$$

Where:

$MS \equiv$ Mean square due to regression.

$$MS = \sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2 / df_m \quad \text{.....(2.9)}$$

$(Y_i - \overline{Y}) \equiv$ The deviation of the i^{th} simulated value from the series mean.

$df_m \equiv$ Degrees of freedom of $[\ \sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2\ ]$

$S^2 \equiv$ Mean sum due to residual variation.

$$S^2 = (SSE / df) \quad \text{.....(2.10)}$$

$$bias = (+/-)(\hat{V} - V) / V * 100 \quad \text{.....(2.11)}$$

where:

$\hat{V} \equiv$ The total estimated volume.

$V \equiv$ The actual volume. Positive (+) values indicate overestimation, and negative (-) values indicate underestimation.

$$\%MF = (+/-)(Ms - Ma) / Ma * 100 \quad \text{.....(2.12)}$$

where:

$Ms \equiv$ The simulated maximum (peak) flow.

$Ma \equiv$ The actual maximum flow. Positive (+) values indicate overestimation, and negative (-) values indicate underestimation.

$$RMSE = \sqrt{(\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 / n} \quad \text{.....(2.13)}$$

2.4.2 **Model Verification**: -

Because the results of any calibration process are conditional on several factors – the calibration data, the objective function, and the optimization procedure – it is a good practice to conduct a verification test. The usual method is to test the performance of calibrated model on a selected portion of the data that were not used for calibration.

These verification data should contain hydrologic behavior of the kind under which the model is expected to provide streamflow estimates or forecasts. Such a test is often referred to as a "split-sample" test – test part of the data record is to detect any biases that may have crept into parameter estimates due to imperfections in the calibration procedure. For example, it is not uncommon for the calibrated model to fit the calibration data very well, but then be found to not perform as well on the validation period – a phenomenon known as "model divergence". When such behavior is observed, it may be necessary to critically examine the entire calibration procedure to determine if certain assumptions are not appropriate or valid and to thereby revise the procedure (Singh, 1997).

2.5 Models Used in the Nile River and its Tributaries: -

In the wake of the devastating flood of 1988, which caused considerable damages in the Khartoum area and northern Sudan, a comprehensive reconstruction program was launched and financed by the World Bank. Besides repair of damages and rehabilitation of infrastructures, the reconstruction program comprised component that would enable Sudan to cope with future floods: the establishment of a flood early warning system (FEWS). FEWS was established in 1991-1992, for the Blue Nile and Atbara River. A Flood Early Warning Center was established at the Ministry of Irrigation and Water Resources in Khartoum, receiving real-time data from METEOSAT as well as from monitoring points with Ethiopia down to Dongla, some 320 km south of the border with Egypt (Ogink et al, 1994).

Besides flood forecasting, the FEWS has also proven to be an efficient and reliable system for water management, especially for the day-to-day reservoir management. In addition, the system provides necessary information to improve other water management aspects, such as storage for irrigation and hydropower generation, as well as reservoir sedimentation.

The rainfall solely generates Nile floods in the Sudan over the upper catchments of the Blue Nile and Atbara River in Ethiopia and Eritrea. Rainfall in the Sudan hardly contributes to the Nile flow. The lead-time between rainfall and the river flow at various spots along the Nile River is shown in Figure (2.4). It shows that by accounting for the rainfall-runoff process in the upper catchments extra gains in the lead-time of one to two days (Atbara-Setit) and of three days (Blue Nile) are obtained relative to river flows forecasting starting off from the border stations. Therefore, the Nile FEWS covers the Blue Nile and Atbara basins fully. The White Nile, although contributing little to the Nile floods, is included in the FEWS for its lower part downstream of Malakal in order to incorporate possible water management options through the Jebel Aulia reservoir. The FEWS extends as far as Dongola on the main Nile.

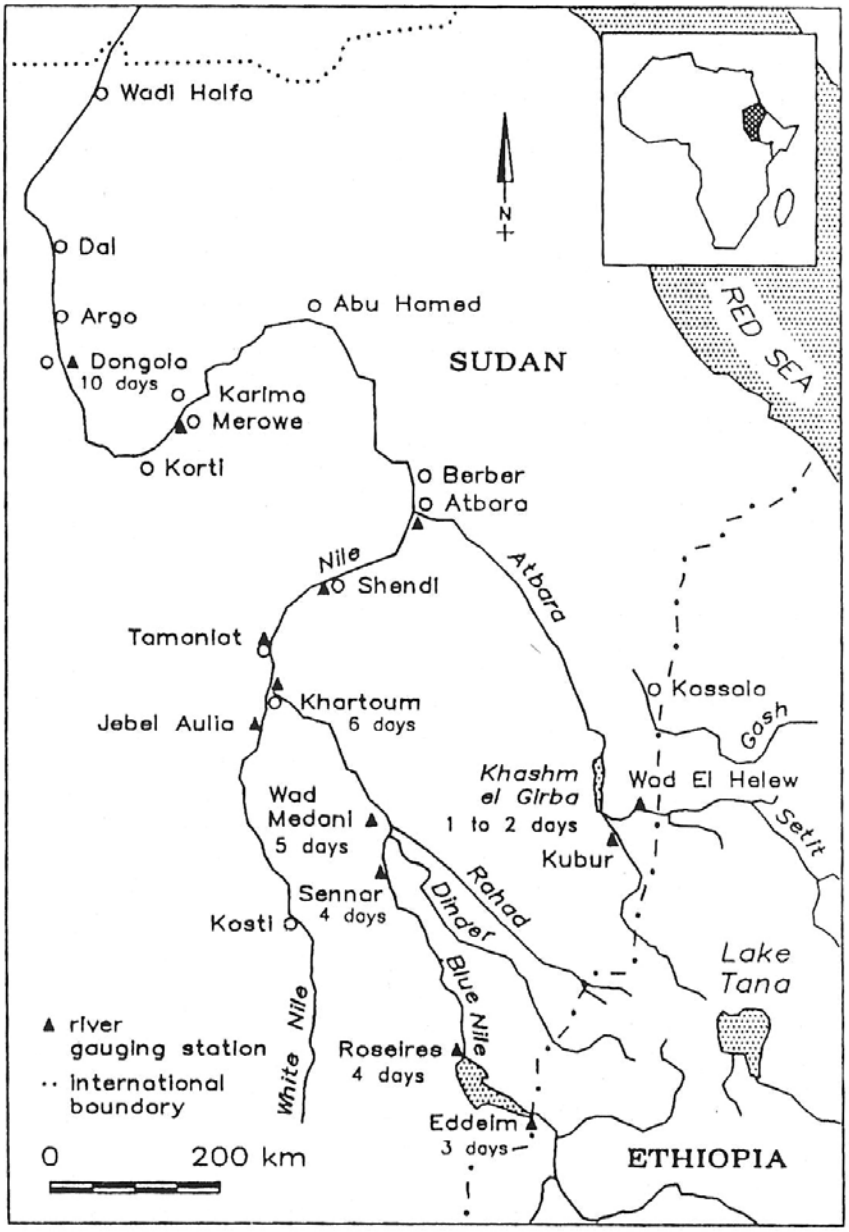


Figure (2.5): The area covered by FEWS with the lead-time between rainfall and the river flow at various spots along the Nile River (Ogink et al, 1994).

FEWS consists of three main components:

1- A Primary Data User Station (PDUS) with relevant software for receiving and processing METEOSAT thermal infra-red (TIR) images on a half-hour basis. These data

are used for estimation of daily rainfall quantities from cold cloud duration and coverage data over the cathments of the Blue Nile and Atbara rivers.

2- A communication system for real-time transmission of water levels in the Blue Nile, Atbara and Main Nile in the Sudan to the Flood Warning Center in Khartoum.

3- A computerized Flood Forecasting System, consist of rainfall-runoff and flow routing models and a temporary database with an appropriate user-interface, allowing smooth and rapid data processing and forecasting.

The basis for the mathematical models in the FEWS forms DELFT HYDRAULICS' SAMFIL rainfall-runoff (based on the well Known Sacramento model) and NETFIL flow routing models respectively. The upper parts of the Blue Nile, Atbara and Setit rivers are covered by the rainfall-runoff models, which produced inflow to the river system at the Ethiopian-Sudanese borders. From their onward the river routing models takes over to route the flow through the main river system reservoirs. The models are equipped with a data assimilation algorithm (Kalman Filter), which combines calculated results on water levels and flows with actual observations, received daily via the HF-radio communication system. This allows the state of models to be updated daily in the line with actual observed state of the rivers, thus create the best possible starting-point for the next forecasting run. Each day during the flood season forecasts is for a period of ten days. This is approximately the lead-time between rainfall events over Ethiopian and the rise in water level at Dongola.

The FEWS started its operation in the flood season of 1992; the experience of two flood season later (1993&1994) yields the following: the quality of the forecasts in the lower part of the system even several days ahead is very satisfactory as a result of the excellent performance of the flow routing part of the system. In which the error in the water level forecasts at Khartoum and Shendi 3 days ahead was only 0.05 to 0.15 respectively, whereas 5 to 6 days ahead the error were still less than 0.15 to 0.20 m. Further upstream the differences between measurement and forecast grew somewhat due to less accurate

inflow forecasts produced by the rainfall-runoff model. Basically four factors are accountable for this:

1- Inaccuracy in the rainfall estimation because differences in local meteorological conditions appear to affect the rainfall-CCD relation.

2- Limited quality of short and medium range rainfall forecast.

3- Doubtful quality of evaporation estimation. Evaporation is an important component in the catchment water balances, but insufficient climatological and vegetation data from the Ethiopian part of the Blue Nile, Atbara and Setit basins are available for accurate estimates.

4- The absence of any runoff data from inside the upper catchments, and hence lacking knowledge on their drainage characteristics have led to a strong lumped in the rainfall-runoff modeling for these catchments. Due to this any physical interpretation to the model parameters has become meaningless: the models have become black boxes. Sofar, a more distributed approach, to account better for the spatial variation in the rainfall alone, but without any knowledge about the drainage characteristics of the sub-catchments, did not prove to be effective (Ogink et al, 1994).

Another project to Monitor, Forecast and Simulate (MFS) flows along the Nile River was initiated and funded in 1991 by the U.S. Agency for International Development (AID). The executing agency to the project was the United Nations Food and Agricultural Organization (FAO). The prime contractor in the project is the National Oceanic and Atmospheric Administration National Weather Service (NOAA). The prime purpose of the MFS project is to develop a hydrometeorological forecast system for the Nile River flows into High Aswan Dam. In additional goal is to regionalize forecast capability so that many of the Nine Nilotic countries could benefit from use of the Nile River System (Attia et al, 1994).

Phase I of the project (September 1990- September1993) involved initial development of a Nile forecast systems (NFS) for the Blue Nile River, since 80% of the runoff of the Nile originates from the Blue Nile system. It was completed with the installation of the latest version of the (NFS) in July 1993. Phase II (September 1993-September 1995) (Karyabwite, 2000), (Attia, et al, 1994) involves significant improvement of the accuracy and simulation of the NFS for the Blue Nile as well as expansion of the system to include the White Nile River system.

The Nile Forecasting System consists of two sub-systems. The Primary Data User System (PDUS) provides the continuous input of METEOSAT satellite data to the NFS, and hydrologic models and software required to produce river and flow forecasting along the Nile River system.

Real-time METEOSAT satellite imagery data received every 30minutes by the PDUS. The NFS then processes the imagery data with the final output being flow hydrographs at designated locations along the Blue Nile River.

The Nile forecast system consists of a hydroclimate data base system, a preprocessor component, a forecast component, a user inerface, and GIS. The computation of data and forecasts is based on the METEOSAT Grid system, which is quasi-rectangular 5.5 km grid scale there are 11,000 grid cells in the Blue Nile Basin each having an area of about 30 km². The Hydroclimate component consist of an observed precipitation time series data base, a climate data base, observed flow and stage data, reservoir data, potential evapotranspiration data, monthly girded precipitation files foe more that 50 years, and a METEOSAT raw imagery data base.

The forecast component consists of distributed hydrological models and software to produce flow and stage hydrographs based on precipitation inputs supplied by the processor. The hydrologic operations consist of: a sub-watershed model to compute surface runoff at grid points and sub-areas, a hillslope routing model to route flow from pixel to a stream, a channel routing model, a stochastic hydrologic model for the White

Nile Basin, an Assimilator to filter rainfall inputs and state variables and a technique to use climatological information to predict long term flow forecasts & uncertainty. The runoff simulation model operates in the Blue Nile and Atbara River basins. The runoff simulation model has three parts: a water balance model developed by Koren and Schaake in 1992 (Attia et al, 1994) which produces surface runoff; a single, none-linear storage equation to route the surface runoff within the cell of the stream, and a single none-linear storage equation for moving water from grid cell to grid cell. The runoff grid cell model has fourteen parameters and four state variables, the state variables include: soil moisture storage, surface moisture storage, overland flow storage, and channel storage. There are two reservoirs in the Blue Nile basin where reservoir routing techniques are applied. These reservoirs are Sennar and Roseires.

The first operational forecast for the Nile was issued in July 1993. The forecast predicted 17% above normal volume inflows into Aswan Dam for the period July-November 1993. This forecast assumed near normal precipitation conditions during the forecast months (Attia et al, 1994), (Karyabwite, 2000).

Chapter 3

THE FUNDAMENTALS OF THE MODELS USED DURING THE STUDY

3.1 The Artificial Neural Networks: -

For many centuries, one of the goals of human kind has been to develop machines. These machines were envisioned as performing all cumbersome and tedious tasks so that a more fruitful life might be enjoyed. The era of machine making began with discovery of simple machines such as lever, wheel and pulley. Many equally congenial inventions followed thereafter. Nowadays engineers and scientists are trying to develop intelligent machines. Artificial neural systems are present-day examples of such machines that have great potential to further improve the quality of our life.

The recent resurgence of interest in neural networks has its roots in the recognition that the brain performs computations in a different manner than do conventional digital computers. Computers are extremely fast and precise at executing sequences of instructions that have been formulated for them. A human information processing system is composed of neurons switching at speed about a million times slower than computer gates. Yet human are more efficient than computers at computationally complex tasks such as speech understanding. Moreover, not only humans, but even animals, can process visual information better than the fastest computers.

People and animals are much better and faster at recognizing images than most advanced computers. Although computers outperform biological and artificial neural systems for tasks based on precise and fast arithmetic operations, artificial neural systems represent the promising new generation of information processing networks. Advances have been made in applying such systems for problems found intractable or difficult for traditional computation. Neural networks can supplement the enormous processing power of the von Neumann digital computer with the ability to make sensible decisions and to learn by ordinary experience, as human do.

A neural network's ability to perform computations is based on the hope that they can reproduce some of the flexibility and power of the human brain by artificial means. Network computation is performed by a dense mesh of computing nodes and connections. They operate collectively and simultaneously on most or all data and inputs. The basic processing elements of neural networks are called artificial neurons, or simply neurons. Often simply they were called them nodes. Neurons perform as a summing and nonlinear mapping junctions. In some cases they can be considered as threshold units that fire when their total input exceeds certain bias levels. Neurons usually operate in parallel and more are configured in regular architectures. They are often organized in layers, and feedback connections both within the layer and toward adjacent layers are allowed. Each connection strength is expressed by a numerical value called a weight, which can be modified. The effort is not only called as neural networks, but also neurocomputing, network computation, connectionism, parallel distributed processing, layered adaptive systems, self-organizing networks, or neuromorphic systems or networks.

Artificial neural systems function as parallel distributed computing networks. Their most basic characteristic is their architecture. Only some of the networks provide instantaneous responses. Other networks need time to respond and are characterized by their time-domain behavior, which refers to as dynamics. Neural networks also differ from each other in their learning modes. There are a variety of learning rules that establish when and how the connecting weights change. Finally, networks exhibit different speeds and efficiency of learning. As a result, they also differ in their ability to accurately respond to the cues presented at the input.

In contrast to conventional computers, which are programmed to perform specific tasks, most neural networks must be taught, or trained. They learn new associations, new patterns, and new functional dependencies. Learning corresponds to parameter changes. Learning rules and algorithm used for experimental training of networks replace the programming required for conventional computers. Neural network users do not specify an algorithm to be executed by each computing node, as would programmers of a more

traditional machine. Instead, they select what in their view is the best architecture, specify the characteristics of the neurons and initial weight, and choose the training mode for the network. Appropriate inputs are then applied to the network so that it can acquire knowledge from the environment. As a result of such explore, the network assimilates the information that can later be recalled by the user. Finally, scientist and technologist are interested in opportunities that are opened by the massively parallel computational networks in the area of artificial intelligence, computational theory, modeling and simulation, and others (Zurada, 1995).

3.1.1 History of Artificial Neural Systems Development: -

Artificial neural systems development has an interested history. Since it is not possible to cover this history in depth in a short paragraph, only major achievements are mentioned. The year 1943 is often considered the initial year in the development of artificial neural systems. In 1943 McCulloch and Pitts outlined the first formal elements to perform logic operations, and thus it could function as an arithmetic-logic computing element. The implementation of its compact electronic model, however, was not technologically feasible during the era of bulky vacuum tubes. However, the McCulloch and Pitts' neuron model laid the groundwork for future development.

In 1949 Donald Hebb first proposed a learning scheme for updating neuron's connections which were refered to as the Hebbian learning rule. He stated that the information can be stored in connections, and postulated the learning technique that had a profound impact on future developments in this field.

During the 1950s, the first neurocomputers were built and tested by Minsky in 1954. They adapted connections automatically. During this stage, Frank Rosenblatt invented the neuron-like element called a perceptron in 1958. It was a trainable machine capable of learning to classify certain patterns by modifying connections to the threshold elements. The idea caught the imagination of engineers and scientist and laid the groundwork for the basic machine learning algorithms that are still used to day. In the

early 1960s a device called ADALINE (for ADAptive LINear combiner) was introduced, and a new, powerful learning rule called the Widrow-Hoff learning rule was developed by Bernard Widrow and Marician Hoff in 1962 (Zurada, 1995). The rule minimized the summed square error during training involving pattern classification.

Despite success and enthusiasm of the early and mid-1960s, the existing machine learning theorems of that time were too week to support more complex computational problems. Although the bottlenecks were exactly identified in Nilsson's work and the neural network architectures called layered networks were also known, no efficient learning schemes existed at that time that would circumvent the formidable obstacles. Neural network research entered into the stagnation phase. Another reason that contributed to this research slowdown at that time was the relatively modest computational resources available at that time.

The final episode of this era was the publication of a book by Marvin Minsky and Seymour Papert in 1969 that gave more doubt as to the layered learning networks potential. The stated limitations of the perceptron-class networks were made public; however the challenge was not answered until the mid-1980s (Zurada, 1995).

During the period from 1982 until 1986, several seminal publications were published that significantly furthered the potential of neural networks. The era of renaissance started with John Hopfield in 1986 introducing a recurrent neural network architecture for associative memories. His paper formulated computational properties of a fully connected network of units.

Another revitalization of the field came from the publication in 1986 of two volumes on parallel distributed processing, edited by James McClellel and David Rumelhart in 1986. The new learning rules and another concepts introduced in this work have removed one of the most essential network training barriers that grounded the mainstream efforts of the mid-1960s (Zurada, 1995).

3.1.2 Biological Neurons and their Artificial Models: -

A human brain consists of approximately 10^{11} computing elements called neurons. They communicate through a connection network of axons and synapses having a density of approximately 10^{14} synapses per neuron. Our hypothesis regarding the modeling of the neural nervous system is that neurons communicate with each other by means of electrical impulses (Arbib 1987). The neurons operate in a chemical environment that is even more important in terms of actual brain behavior. Thus the brain can be considered as a densely connected electrical switching network conditioned largely by the biochemical processes. The vast neural network has an elaborate structure with very complex interconnections. The input to the network is provided by sensor connectors. Receptors deliver stimuli both from within the body, as well as from sense organs when the stimuli originate in the external world. The stimuli are in the form of electrical impulses that convey the information into the network of neurons. As a result of information processing in the central nervous systems, the effectors are controlled and give human responses in the form of diverse actions. There fore there is a three-stage system, consisting of receptors, neural network, and effectors, in control of the organism and its actions. A lucid, although rather approximation idea, about the information links in the nervous system is shown in Figure (3.1). As it can be seen from the figure, the information is processed, evaluated, and compared with the stored information in the central nervous system. When necessary, commands are generated there and transmitted to the motor organs. Notice that motors are monitored in the central nervous system by feedback links that verify their action. Both internal and external feedbacks control the implementation of commands. As can be seen, the overall nervous system structure has many of the characteristics of closed-loop control system.

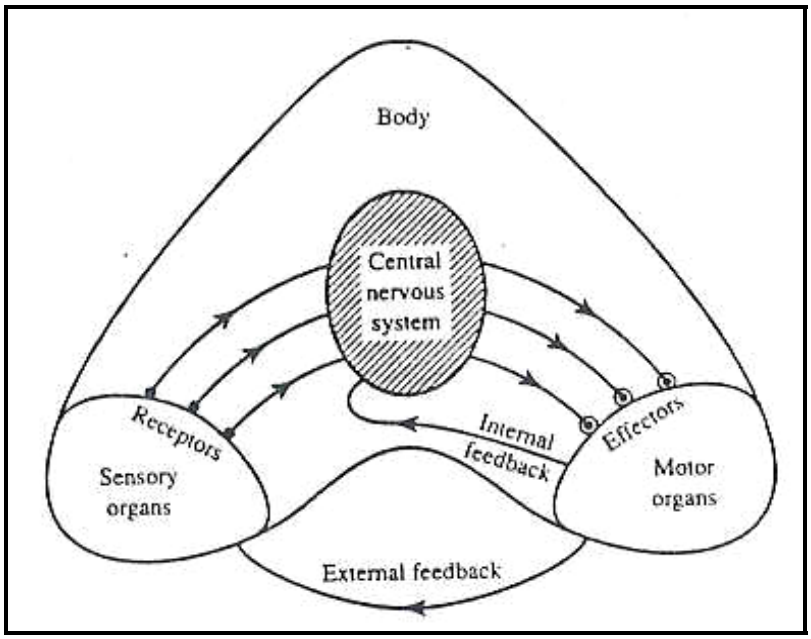


Figure (3.1): Information flow in nervous system (Zurada, 1995)

3.1.2.1 *Biological Neuron*: -

The elementary nerve cell, called a neuron, is the fundamental building block of the biological neural network. Its schematic diagram is shown in Figure (3.2). A typical cell has three major regions: the cell body, which is also called the soma, the axon, and the dendrites. Dendrites form a dendritic tree, which is a very fine bush of thin fibers around the neuron's body. Dendrites receive information from neurons through axons along fibers that serve as transmission lines. An axon is a long cylindrical connection that carries impulses from the neuron. The end part of an axon splits into a fine arborization. Each branch of it terminates in a small end bulb almost touching the dendrites of neighboring neurons. The axon-dendrite contact organ is called synapse. The synapse is where the neuron introduced its signal to the neighboring neuron. The signal reaching a synapse and received by dendrites are electrical impulses. The interneuronal transmission is sometimes electrical but is usually effected by the release of chemical transmitters at the synapse. Thus, terminal boutons generate the chemical that affects the receiving neuron. The receiving either generate an impulse to its axon, or produces no response.

The neuron is able to respond to the total of its inputs aggregated within a short-time interval called the period of latent summation. The neuron's response is generated if the total potential of its membrane reaches a certain level. The membrane can be considered as a shell, which aggregates the magnitude of the incoming signals over some duration. Specifically, the neuron generates a pulse response and sends it to its axon only if the conditions necessary for firing are fulfilled.

When the conditions necessary for the firing of a neuron is considered, incoming impulses can be excitatory if they cause the firing, or inhibitory if they hinder the firing of the response. A more precise condition for firing is that the excitation should exceed the inhibition by the amount called the threshold of the neuron, typically a value of about 40mV (Arbib 1987). Since a synaptic connection causes the excitatory or inhibitory reactions of the receiving neuron, it is practical to assign positive and negative unity weight values respectively, to such connections. This allows the neuron's firing condition to be formulated. The neurons firs when the total of the weights to receive impulses exceeds the threshold value during the latent summation period.

The incoming impulses to a neuron can only be generated by neighboring neurons and by the neuron itself. Usually, a certain number of incoming impulses are required to make a target cell fire. Impulses that are closely spaced in time and arrive synchronously are more likely to cause the neuron to fir. As mentioned before, observations have been made that biological networks perform temporal integration and summation of incoming signals. The resulting spatio-temporal processing performed by natural neural networks is a complex process and much less structured than digital computation. The neural impulses are not synchronized in time as opposed to the synchronous discipline of digital computation. The characteristic feature of the biological neuron is that the signals generated do not differ significantly in magnitude: the signal in the nerve fiber is either absent or has the maximum value. In other words, information is transmitted between the nerve cells by means of binary signals.

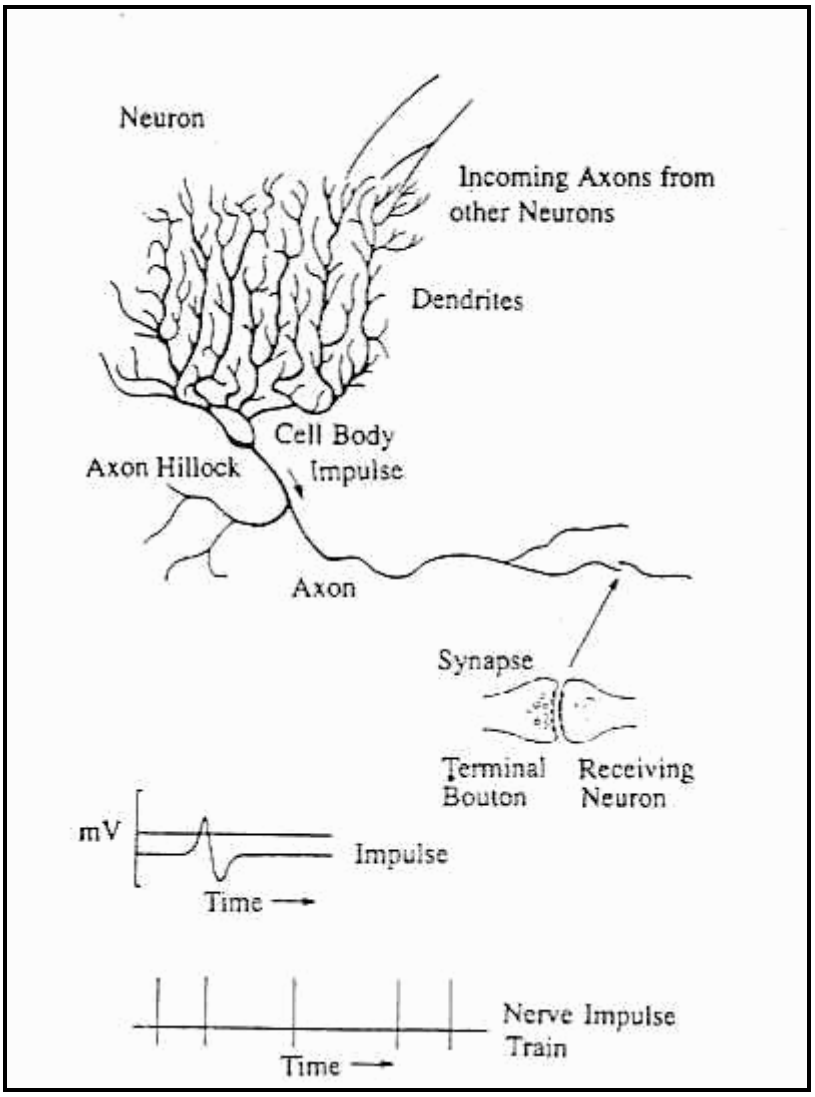


Figure (3.2): Schematic diagram of a neuron and a sample of pulse train (Zurada, 1995).

After carrying a pulse, an axon fiber is in a state of complete non-excitability for a certain time called the refractory period. For this time interval the nerve does not conduct any signals, regardless of intensity of excitation. Thus, the time scale can be divided into consecutive intervals, each equal to the length of the refractory period. This will enable a discrete-time description of the neurons' performance in terms of their states at a discrete time instance. For examples, it is easy to specify which neurons will fire at the instance $k+1$ based on the excitation conditions at the instant k. The neuron will be excited at the present instant if the number of excited excitatory synapses exceeds the number of

excited inhibitory synapses at the previous instant by at least the number T, where T is the neuron's threshold value.

The time units for modeling biological neurons can be taken to be of the order a millisecond. However, the refractory period is not uniform over the cells. Also, there are different types of neurons and different ways in which they connect. Thus, the picture of real phenomena in the biological neural network becomes even more involved and dealing with a dense network of interconnected neurons that release asynchronous signals. The signals are then fed forward to the other neurons within the spatial neighborhood but also back to the generating neurons.

3.1.2.2 *Neuron Modeling for Artificial Neural Systems*: -

3.1.2.2.1 *McCulloch-Pitts Neuron Model*: -

The first formal definition of a synthetic neuron model based on the highly simplified considerations of the biological model, which described in the preceding section, was formulated by McCulloch and Pitts 1943. This model of the neuron is shown in Figure (3.3). The inputs x_i, for I= 1,2… n, are 0 or 1, depending on the absence or presence of the input impulse at instant k. The neuron's output signal is denoted as o. the firing rule for this model is defined as follows:

$$O^{k+1} = \begin{cases} 1 & if \sum_{i=1}^{n} w_i x_i^k \geq T \\ 0 & if \sum_{i=1}^{n} w_i x_i^k < T \end{cases} \quad \text{.....(3.1)}$$

Where superscript k =0,1,2… denotes the discrete-time instant, and w_i is the multiplicative weight connecting the $i^{'}th$ input with the neuron's membrane. Note that w_i =+1 for excitatory synapses, w_i = - 1 for inhibitory synapses for this model, and t is

the neuron's threshold value, which needs to be exceeded by the weighted sum of signals for the neuron to fire.

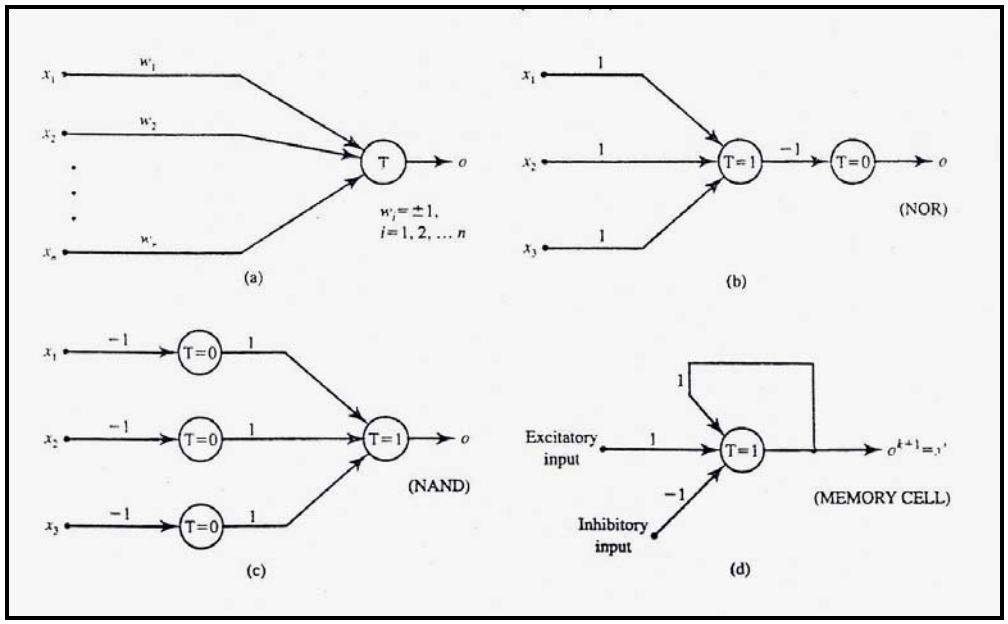


Figure (3.3): McCulloch-Pitts model neuron and elementary logic networks: (a) model neuron, (b) NOR Gate, (c) NAND Gate, and (d) Memory Cell.

3.1.2.2.2 *The Artificial Neuron*: -

Every neuron model consists of a processing element with synaptic input connections and a single output. The signal flow of neuron inputs x_i is considered to be unidirectional as indicated by arrows, as a neuron's output signal flow. A general neuron symbol is shown in Figure (3.4). This symbolic representation shows a set of weights and the neuron's processing unit, or node. The neuron output signal is given by the following relationship:

$$O = f(W^t X), \text{ Or} \quad \text{.....(3.1a)}$$

$$O = f(\sum_{i=1}^{n} w_i x_i) \quad \text{.....(3.1b)}$$

Where W is the weight vector defined as

$$W \triangleq [w_1 w_2 w_3 \cdots w_n]^t \quad \text{.....(3.2)}$$

And X is the input vector

$$X \triangleq [x_1 x_2 x_3 \cdots x_n]^t \quad \text{.....(3.3)}$$

(All vectors defined in this chapter are column vectors; superscript t denotes a transposition.) The function $f(w^t x)$ is often referred to as an activation function. Its domain is the set of activation values, net of the neuron model; thus this function was often used as $f(net)$. The variable net is defined as a scalar product of the weight and input vector

$$net \triangleq w^t x \quad \text{.....(3.4)}$$

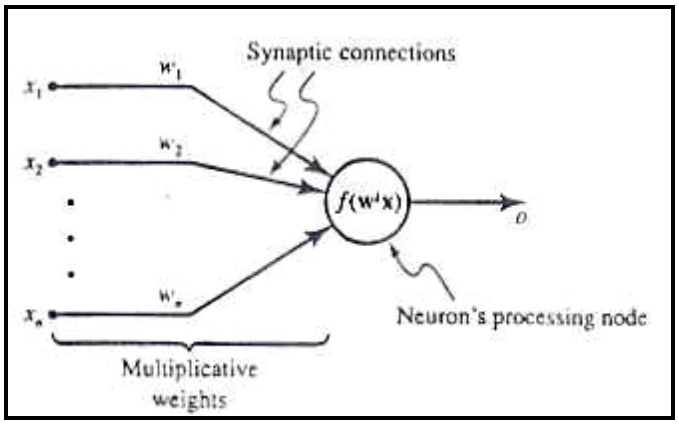


Figure (3.4): General symbol of neuron consisting of processing node and systematic connections.

The argument of the activation function, the variable net, is an analog of the biological neuron's membrane potential. Note that temporarily the threshold value (sometimes known as thresholding bias) is not explicitly used in Equations (3.1) & (3.4), but this is only for notational convenience. It has to be momentarily assumed that the modeled

neuron has $n-1$ actual synaptic connections that come from actual variable inputs $x_1, x_2, \cdots x_{n-1}$. Also, it has to be assumed that $x_n = -1$ and $x_n = t$. Since threshold plays an important role for some models; sometimes it will be needed to extract explicitly the threshold as separate neuron model parameters (with a bias weight) see Figure (3.5). Sometimes an auxiliary input is included as an input to the perceptron to serve as a thresholding bias.

The general neuron symbol, shown in Figure (3.4) and descried by Equations (3.1) and (3.4), are commonly used in neural network literature. However, different artificial neural network classes make use of different definitions of $f(net)$. Also, even within the same class of networks, the neurons are sometimes considered to perform differently during different phases of network operation. Therefore, it is pedagogically sound to replace, whenever needed, the general neuron model symbol from Figure (3.4) with a specific $f(net)$ and a specific neuron model. The model validity will then usually be restricted to a particular class of network. Two main models introduced below are often used.

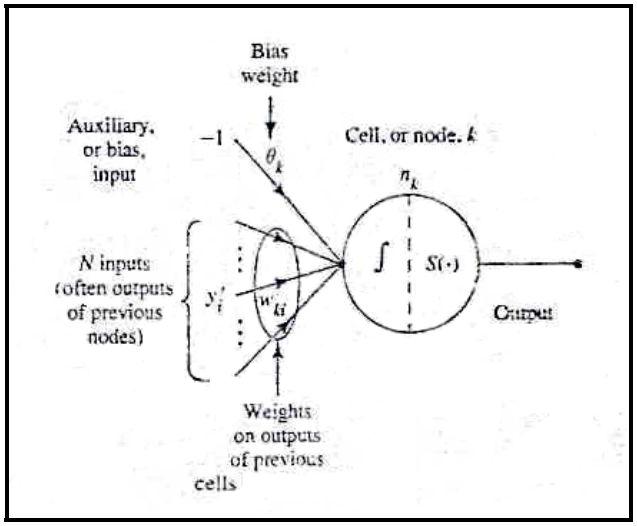


Figure (3.5): Formal labeling of artificial neuron.

Observe from Equation (3.1) that the neuron as a processing node performs the operation of summation of its weighted inputs, or the scalar product computation to obtain net.

Subsequently, it performs the nonlinear operation $f(net)$ through its activation function. Typical activation functions used are

$$f(net) \overset{\Delta}{=} \frac{2}{1+\exp(-\lambda net)} - 1 \quad \text{..... (3.5a)}$$

And

$$f(net) \overset{\Delta}{=} \operatorname{sgn}(net) = \begin{cases} +1, & net > 0 \\ -1, & net < 0 \end{cases} \quad \text{..... (3.5b)}$$

Where $\lambda > 0$ in (3.5a) is proportional to the neuron gain determining the steepness of the continuous function $f(net)$ near $net = 0$. The continuous activation function is shown in Figure (3.6) for various λ. Note that as $\lambda \to \infty$, the limit of the continuous function becomes the $\operatorname{sgn}(net)$ function defined by Equation (3.5b). Activation functions defined by Equations (3.5a) & (3.5b) are called bipolar continuous and bipolar binary functions, respectively. The word "bipolar" is used to point out that both positive and negative responses of neurons are produced for this definition of the activation function.

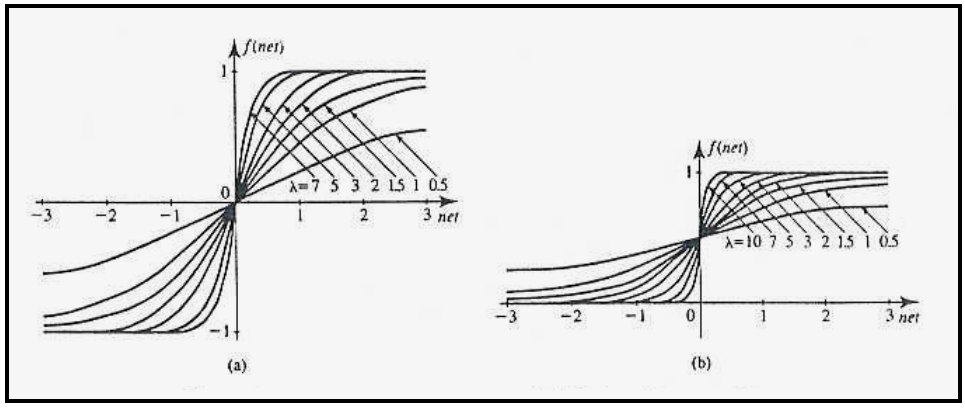


Figure (3.6): Activation functions of a neuron: (a) bipolar continuous and (b) unipolar continuous.

By shifting and scaling the bipolar activation function defined by Equation (3.5), unipolar continuous and unipolar binary activation functions can be obtained, respectively, as

$$f(net) \stackrel{\Delta}{=} \frac{1}{1+\exp(-\lambda net)} \quad \text{..... (3.6a)}$$

and

$$f(net) \stackrel{\Delta}{=} \begin{cases} 1, & net > 0 \\ 0, & net < 0 \end{cases} \quad \text{..... (3.6b)}$$

Function (3.6a) is shown in Figure (3.6). Again, the unipolar function is the limit of $f(net)$ in Equation (3.4) when $\lambda \rightarrow \infty$. The soft-limiting activation functions shown by Equations (3.3) & (3.4) are often called sigmoidal characteristics, as opposed to the hard limiting activation function given in Equations (3.3) & (3.4). Hard limiting activation functions describe the discrete neuron model.

Neuron outputs are either discrete or continuous. Given a layer of m neurons, their output values $o_1 o_2 \cdots o_m$ can be arranged in a layer's output vector:

$$O \stackrel{\Delta}{=} [o_1 o_2 \cdots o_m]^t \quad \text{..... (3.7)}$$

Where o_i act as output signal of the i^{th} neuron.

3.1.2.2.3 *Types of Activation Functions*: -

The hidden layers produce outputs based upon the sum of weighted values passed to them. So does the output layer. The way they produce their outputs is by applying an "activation" function to the sum of the weighted values. The activation function, also called the squashing function, maps this sum into the output value, which is then "fired" on to the next layer. Although the logistic function is the most popular, there are other functions that may be used.

Here are some types of these functions:

Logistic Function: -

$$f(x) = \frac{1}{(1+\exp(-x))} \quad \text{.....(3.8)}$$

Linear Function: -

$$f(x) = x \quad \text{.....(3.9)}$$

Tanh Function (the hyperbolic tangent function): -

$$f(x) = \tanh(x) \quad \text{.....(3.10)}$$

Tanh15 Function: -

$$f(x) = \tanh(1.5x) \quad \text{.....(3.11)}$$

Sine Function: -

$$f(x) = \sin(x) \quad \text{.....(3.12)}$$

Symmetric Logistic Function

$$f(x) = \frac{2}{(1+\exp(-x))} - 1 \quad \text{.....(3.13)}$$

Gaussian Function: -

$$f(x) = \exp(-x^2) \quad \text{.....(3.14)}$$

Gaussian-Complement Function: -

$$f(x) = 1 - \exp(-x^2) \quad \text{.....(3.15)}$$

The following are guidelines for when to use the various functions, but like so much in neural networks, there will always be exceptions to these guidelines:

Logistic (Sigmoid logistic) function (Figure (3.7)), it was found that this function useful for most neural network applications. It maps values into the (0, 1) range. Always use this function when the outputs are categories.

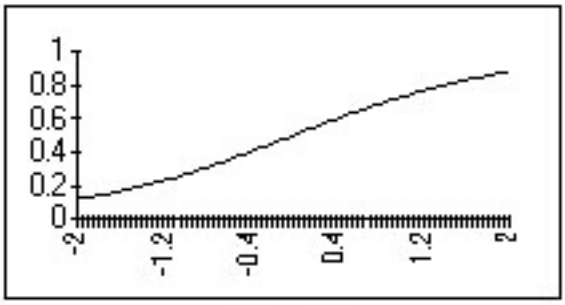


Figure (3.7): The Logistic Function.

Linear function (Figure (3.8)), the use of this function should generally be limited to the output slab. It is useful for problems where the output is a continuous variable, as opposed to several outputs, which represent categories. Stick to the logistic for categories. Although the linear function detracts from the power of the network somewhat, it sometimes prevents the network from producing outputs with more error near the minimum or maximum of the output scale. In other words the results may be more consistent throughout the scale. If it is used, stick to smaller learning rates, momentums, and initial weight sizes. Otherwise, the network may produce larger and larger errors and weights and hence not ever lower the error. The linear activation function is often ineffective for the same reason if there are a large number of connections coming to the output layer because the total weight sum generated will be high.

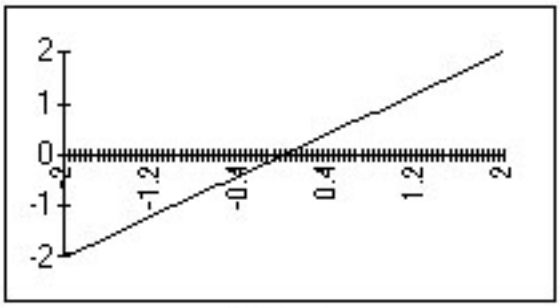


Figure (3.8): The Linear Function.

Tanh (hyperbolic tangent) function (Figure (3.9)), it is sometimes better for continuous valued outputs, however, especially if the linear function is used on the output layer. If it

is used it in the first hidden layer, the inputs should be scaled into [-1, 1] instead of [0, 1]. Good results might be obtained when using the hyperbolic tangent in the hidden layer of a 3-layer network, and using the logistic or the linear function on the output layer.

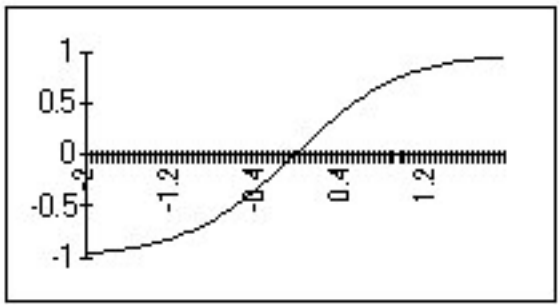


Figure (3.9): The Hyperbolic Tangent (tanh) Function.

Tanh15 (hyperbolic tangent 1.5) function (Figure (3.10)), there has been at least one technical paper wherein it was strongly proposed that tanh(1.5x) is much better than tanh(x).

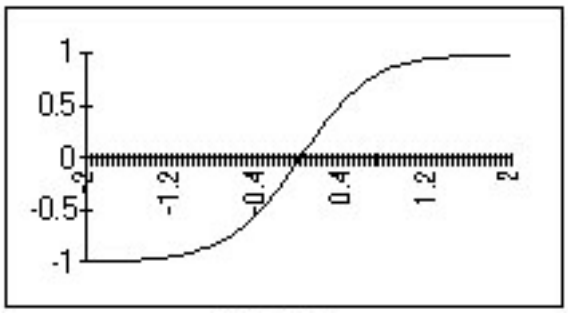


Figure (3.10): The Hyperbolic Tangent1.5 (tanh15) Function.

Sine function (Figure (3.11)), if it is used in the first hidden layer, the inputs should be scaled into [-1, 1] instead of [0,1]. If it is used on the output layer, the outputs should be scaled to [-1,1] also.

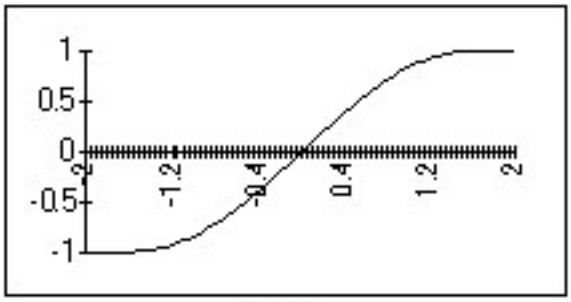


Figure (3.11): The Sine Function.

Symmetric Logistic function (Figure (3.12)), this is like the logistic, except that it maps to (-1, 1) instead of mapping to (0, 1). The different scaling may be better for some problems, and it should be tried. When the outputs are categories, it is preferred to use the symmetric logistic function instead of logistic in the hidden and output layers. In some cases, the network would be trained to a lower error in the training and test sets.

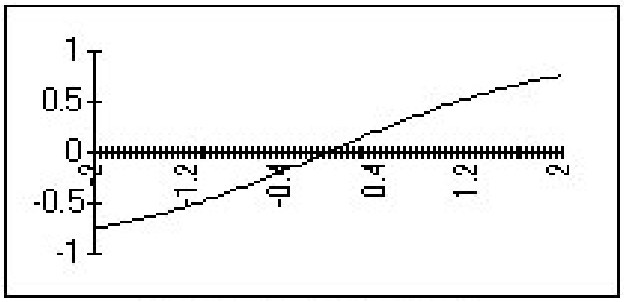


Figure (3.12): The Symmetric Logistic Function.

Gaussian function (Figure (3.13)), this function is unique, because unlike the others, it is not an increasing function. It is the classic bell shaped curve, which maps high values into low ones, and maps mid-range values into high ones. There is not much about its use in the literature, but it has been found that it is very useful in a small set of problems. It is suspected that it brings out meaningful characteristics not found at the extreme ends of the sum of weighted values. This function produces outputs in [0,1]. It was found that an architecture of a three layer network with two slabs in the hidden layer (this is one of the Ward backpropagation networks) has worked very well, by putting a Gaussian

activation function in one slab and Tanh in the second, and then use the logistic activation function in the output layer. For prediction of continuous valued outputs, it is preferred to use a three-layer backpropagation network with a Gaussian in the hidden layer, and a linear activation function in the output layer. For classification problems, it is preferred to use a regular three-layer backpropagation network with a Gaussian in the hidden layer, and a sigmoid logistic function in the output layer.

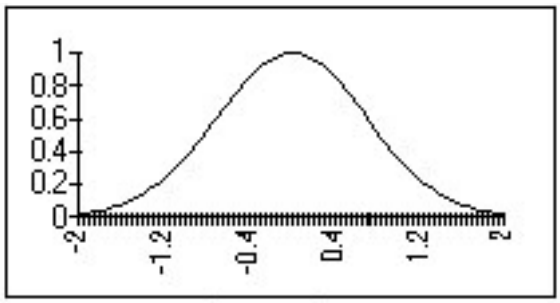


Figure (3.13): The Gaussian Function.

Gaussian Complement function (Figure (3.14)), it was found that this function tends to bring out meaningful characteristics in the extremes of the data.

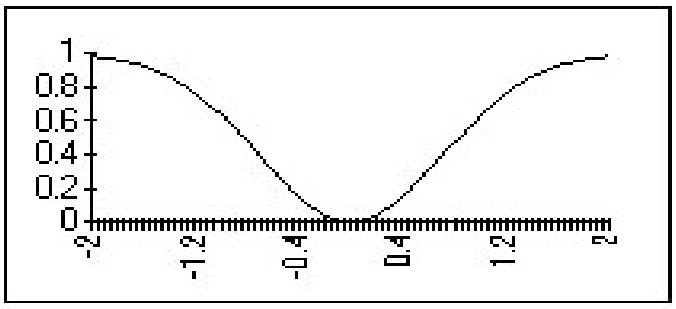


Figure (3.14): The Gaussian Complement Function.

Some of these activation functions may require lower learning rates and momentums than the logistic function requires; in fact these other activation functions may thrive on different learning rates and momentums than the default settings for the logistic function. Also, some sort of carefulness must be taken when a particular function solves the considered problem faster: it should be assured that it also solves it as well, i.e., that the

net generalizes as well. Tanh and sine have steeper slopes than the logistic or symmetric logistic functions, and so they may appear to solve the problem faster. Also note that because some activation functions scale data between 0 and 1 while others scale data between -1 and 1, a different average error for architectures which use different activation functions on the output layer might be obtained (Extracted from NeuroShell 2 software Help, (C) Ward Systems Group, Inc.).

3.1.3 Models of artificial neural network: -

3.1.3.1 *Types of ANN architectures*: -

Having introduced the basic principles of the artificial neuron, it is easy now to begin to connect these unites into networks of cells comprising machines. The resulting ANNs are so named because their topologies consist of "axon-to-dendrite" linkage of the individual cells reminiscent in a primitive way of patterns of biological neurons. The computing power of the resulting network derives from the complex interaction of many simple nonlinear elements that perform their operations in parallel. Many ANN architectures can be conveniently viewed as "layers" of cells, as illustrated in Figure (3.15). A layered structure is one that may be described as follows: A group of N_l cells designated layer 1 receives as their inputs weighted versions of the external inputs of the network. There are No external inputs, one of which may correspond to a bias. The remaining cells in the network (above layer 1) can be grouped into layers 2,3…L, such that cells in layer l receives as inputs weighted outputs of cells in layer $l-1$. The outputs of the final layer l are the external outputs of the network. The term "weighted outputs" loosely to mean some combination of the labels on the connections (weights) with the output of the layer network. This combination rule depends on the operation of the cell. A hidden layer is one containing cells whose outputs cannot be measured directly. According to our framework, each layer is hidden except layer L.

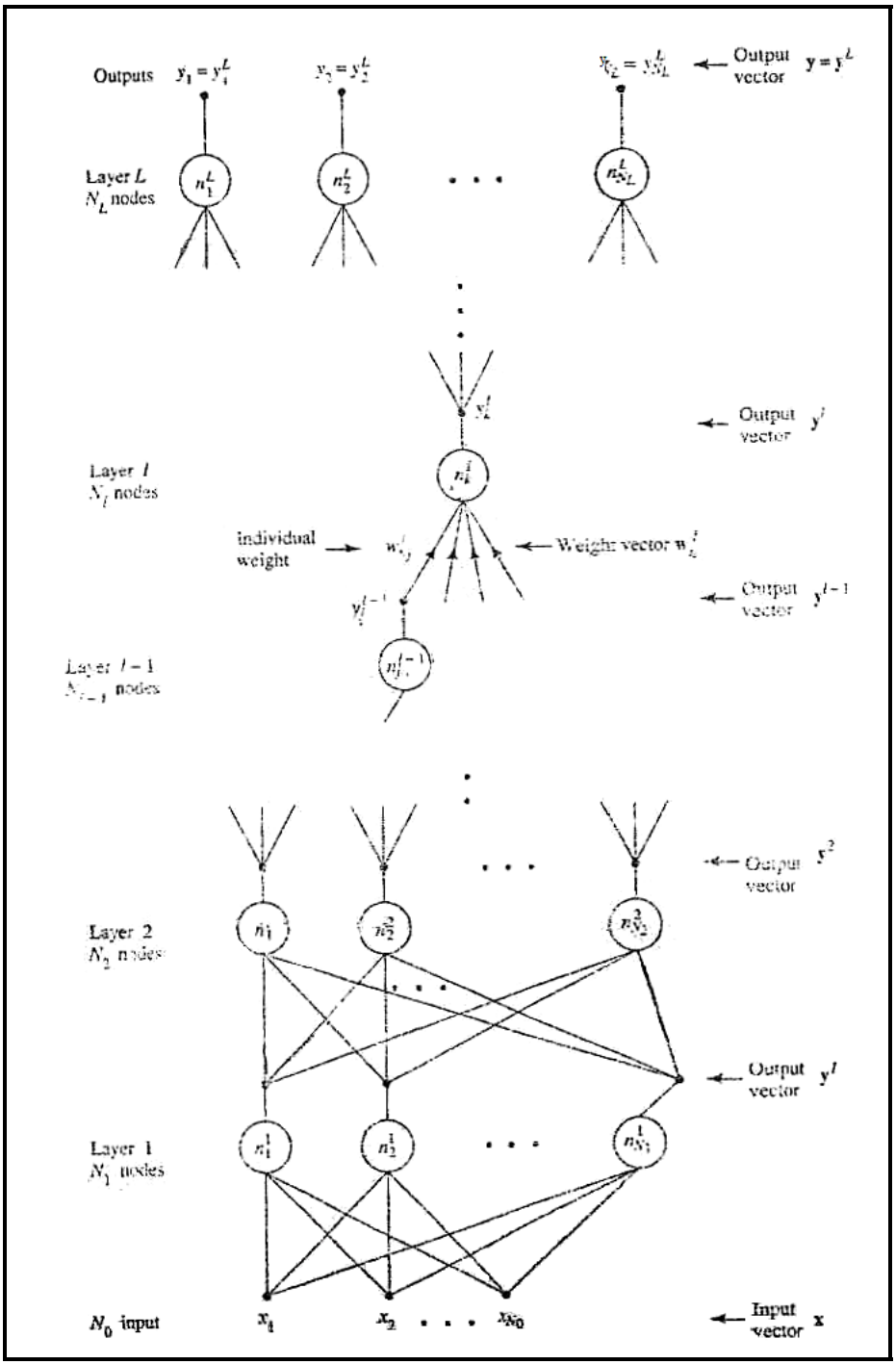


Figure (3.15): A layered artificial neural network.

3.1.3.1.1 *Feedforward ANN*: -

A feedforward, or nonrecurrent, ANN is one for which no cell has a connection path leading from its output back to its input. Layering depends not only on nonrecurrence, but also upon sequential connection. By this it meant that cells in layer l must be connected to cells in layer $l+m$, where m must not only be positive (no feedback) but it must be exactly unity (sequential). A layered ANN can sustain a few feedback or nonsequential connections without losing its basic pattern, but too many such connections erode the fundamental network structure.

When considering an elementary feedforward architecture of m neurons receiving n inputs as shown in Figure (3.16), its output vectors are, respectively:-

$$\mathbf{o} = [o_1 o_2 \cdots o_m]^t \quad \text{.....(3.16a)}$$

$$\mathbf{x} = [x_1 x_2 \cdots x_n]^t \quad \text{.....(3.16b)}$$

Weight w_{ij} connects the i^{th} input. The double subscript convention used for weights is such that the first and second subscripts denote the index of the destination and source, respectively. Thus the activation value for the i^{th} neuron can be written as:

$$net_i = \sum_{j=1}^{n} w_{ij} x_j, \qquad for \quad i = 1,2,...,m \quad \text{.....(3.17)}$$

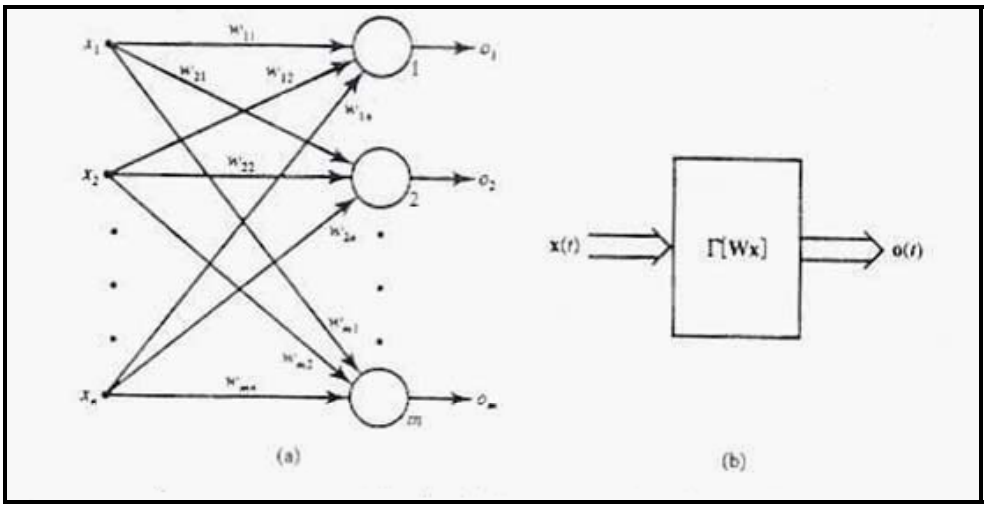


Figure (3.16): Single-layer feedforward network: (a) interconnection scheme and (b) block diagram.

The following nonlinear transformation Equation (3.18) involving the activation function $f(net)$, for $i = 1,2,...,m$, complete the processing of $\mathbf{x}$. The transformation, performed by each of the m neurons in the network, is a strongly nonlinear mapping expressed as:

$$\mathbf{o_i} = f(\mathbf{w_i^t x}), \qquad for \quad i = 1,2,...,m \quad \text{.....(3.18)}$$

Where weight vector w_i contains weights leading toward the i^{th} output node and is defined as follows:

$$\mathbf{w_i} \overset{\Delta}{=} [w_{i1}\, w_{i2} \cdots w_{in}]^t \quad \text{.....(3.19)}$$

Introducing the nonlinear matrix operator (Γ), the mapping of input space x to output space o implement by the network can be expressed as follows

$$\mathbf{o} = \Gamma[\mathbf{Wx}] \quad \text{.....(3.20a)}$$

Where $\mathbf{W}$ is the weight matrix, also known as the connection matrix:

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix} \quad \text{.....(3.20b)}$$

And

$$\Gamma[\cdot] = \begin{bmatrix} f(\cdot) & 0 & \cdots & 0 \\ 0 & f(\cdot) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & f(\cdot) \end{bmatrix} \quad \text{.....(3.20c)}$$

Note that the nonlinear activation functions $f(\cdot)$ on the diagonal of the matrix operator (Γ) operate componentwise on the activation values *net* of each neuron. Each activation value is, in turn, a scalar product of an input with the respective weight vector.

The input and output vector $\mathbf{x}$ and $\mathbf{o}$ are often called input and output patterns, respectively. The mapping of an input pattern into an output patterns as shown in Equation (3.20) is of feedforward and instantaneous type, since it involves no time delay between the input $\mathbf{x}$, and the output $\mathbf{o}$. Thus, Equation (3.20) can be written in the explicit form involving time t as

$$\mathbf{o}(t) = \Gamma[\mathbf{W}\mathbf{x}(t)] \quad \text{.....(3.21)}$$

Figure (3.16) shows the block diagram of the feedforward network. As can be seen, the generic feedforward network is characterized by the lack of feedback. This type of network can be connected in cascade to create a multilayer network. In such a network, the output of a layer is the input to the following layer. Even though the feedforward network has no explicit feedback connection when $\mathbf{x}(t)$ is mapping into $\mathbf{o}(t)$, the output values are often compared with the "teachers" information, which provides the desired output value, and also an error signal can be employed for adapting the network's weights.

3.1.3.1.2 *Feedback ANN*: -

A feedback, or recurrent, ANN is one for which there is a connection path leading from its output back to its input. A feedback network can be obtained from the feedforward shown in Figure (3.16) by connecting the neurons' outputs to their inputs. The result is depicted in Figure (3.17). The essence of closing the feedback loop is to enable control of output o_i through the o_j, for $j = 1,2,\cdots,m$. Such control is especial meaningful if the present output, say $\mathbf{o}(t)$, controls the output at the following instant $\mathbf{o}(t+\Delta)$. The time Δ elapsed between t and $t+\Delta$ is introduced by the delay elements in the feedback loop as shown in Figure (3.17a). Here the time delay d has a symbolic meaning; it is an analogy to the refractory period of an elementary biological neuron model. Using the notation introduced for the feedforward networks, the mapping of o (t) into o $(t+\Delta)$ can now be written as:

$$\mathbf{o}(t+\Delta) = \Gamma[\mathbf{W}\mathbf{o}(t)] \quad \text{.....(3.22)}$$

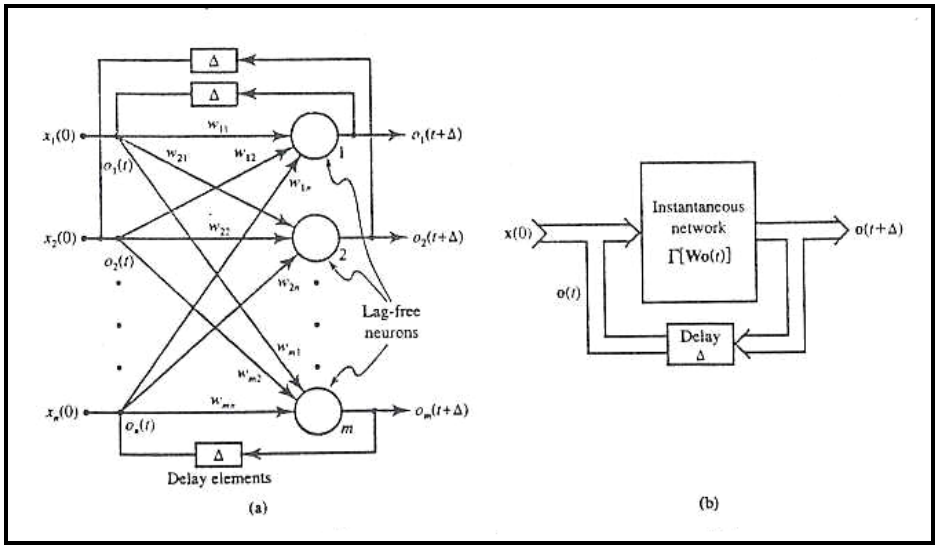


Figure (3.17): Single-layer discrete-time feedback network: (interconnection scheme and (b) block diagram. (Zurada, 1995).

This formula is presented by the block diagram shown in Figure (3.17b). Note that the input $\mathbf{x}(t)$ is only needed to initialize this network so that $\mathbf{o}(0) = \mathbf{x}(0)$. The input is then removed and the system remains autonomous for $t > 0$. Thus a special case of this

feedback configuration can be considered here, such that $x(t) = x(0)$ and no in put is provided to the network thereafter, or for $t > 0$.

3.1.4 Neural Processing: -

The process of computation of **o** (output) for a given **x** (input) performed by the network is known as recall. Recall is the proper processing phase for a neural network, and its objective is to retrieve the information. Recall corresponds to the decoding of the stored content, which may have been encoded in a network previously.

3.1.4.1 *Association*: -

Assume that a set of patterns can be stored in the network. Later, if the network is presented with a pattern similar to a member of the stored set, it may associate the input with the closest stored pattern. The process is called *autoassociation***.** Typically, a degraded input pattern serves as a cue for retrieval of its original form. This is illustrated schematically in Figure (3.18a). The figure shows a distorted square recalling the square encoded.

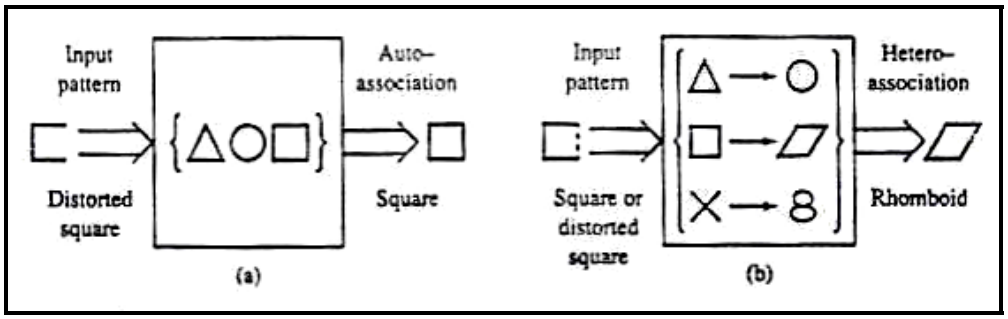


Figure (3.18): Association response: (a) autoassociation and (b) heteroassociation (Zurada, 1995).

Association of input pattern can also be stored in a *heteroassociation* variant. In *heteroassociation* processing, the associations between pairs of patterns are stored. This is schematically shown in Figure (3.18b). A square input pattern presented at the input results in the rhomboid at the output. It can be inferred that the rhomboid and square

constitute one pair of stored patterns. A distorted input pattern may also cause correct *heteroassociation* at the output as shown with dashed line.

3.1.4.2 *Classification*: -

Classification is another form of neural computation. Assume that a set of input patterns is divided into a number of classes, or categories. In response to an input pattern from the set, the classifier is supported to recall the information regarding class membership of the input pattern. Typically, discrete-valued output vectors express classes, and thus output neurons of classifiers would employ binary activation functions. The schematic diagram illustrating the classification response for patterns belonging to three classes is shown in Figure (3.19a).

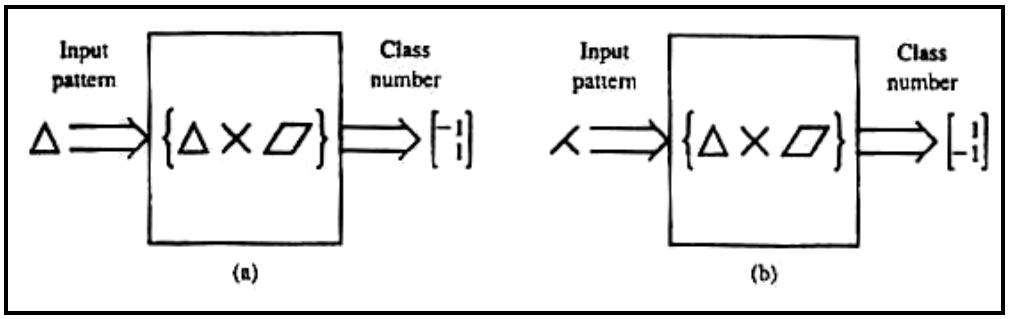


Figure (3.19): Classification response: (a) classification and (b) recognition (Zurada, 1995).

The association is now between the input pattern and the second member of the *heteroassociative* pairs, which is supposed to indicate the input's class number. If the network's desired response is the class number but the input pattern does not exactly correspond to any of the patterns in the set, the process is called *recognition*. When a class membership for one of the patterns in the set is recalled, recognition becomes identically shown in Figure (3.19b). This form of processing is of particular significance when an amount of noise is superimposed on input patterns.

One of the distinct strengths of neural networks is their ability to generalize. The network is said to generalize well when it sensibly interpolates input patters that are new to the

network. Assume that a network has been trained using the data x_1 through x_5, as shown in Figure (3.20). The figure illustrates bad and good generalization examples at points that are new and are between the trained points. Neural networks provide, in many cases, input-output mappings with good generalization capability.

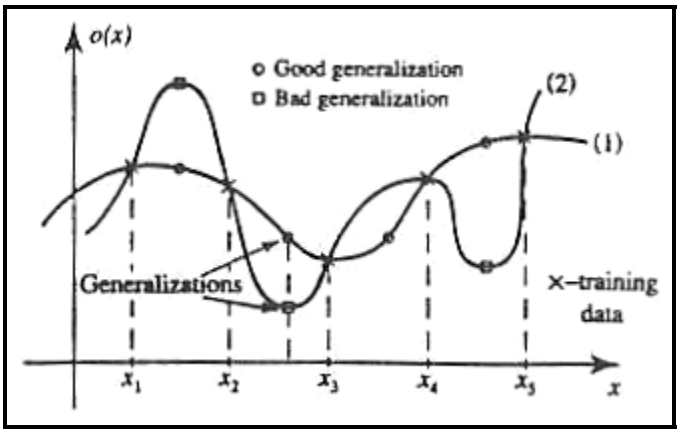


Figure (3.20): Generalization example.

3.1.5 Learning and Adaptation: -

In general, learning is relatively permanent change in behavior brought about by experience. Learning in human beings and animals is an inferred process; cannot see it happenings directly and can assume that it has occurred by observing changes in performance. Learning in neural networks is amore direct process, and typically can capture each learning step in a direct cause-effect relationship. To perform any of the processing tasks, neural network learning of an input-output mapping from a set of examples is needed. Designing an associator or classifier can be based on learning a relationship that transforms input into outputs given a set of examples of input-output pairs.

3.1.5.1 *Learning as Approximation or Equilibria Encoding*: -

Approximation theory focuses on approximating a continuous, multivariable function $h(\mathbf{x})$ by another function $H(\mathbf{w}, \mathbf{x})$, where $\mathbf{x} = [x_1 x_2 \cdots x_m]^t$ is the input vector,

and $\mathbf{w} = [w_1 w_2 \cdots w_m]^t$ is a parameter (weight) vector. In the approach below, a class of neural networks will be considered as systems that can learn approximation of relationships. The learning task is to find $\mathbf{w}$ that provides the best possible approximation of $h(\mathbf{x})$ based on the set of training examples $\{\mathbf{x}\}$. An important choice that needs to be made is which approximation function $H(\mathbf{w}, \mathbf{x})$ to use. An ill chosen, unsmooth, approximation function example is shown in Figure (3.20) as curve (2). Even with the best choice of parameters for an ill chosen function, the approximation is inaccurate between successive data points. The choice of function $H(\mathbf{w}, \mathbf{x})$ in order to represent $h(\mathbf{x})$ is called a representation problem. Once $H(\mathbf{w}, \mathbf{x})$ has been chosen, the network learning algorithm is applied for finding optimal parameters $\mathbf{w}$. A more precise formulation of the learning problem can be stated as calculated involving $\mathbf{w}^*$ such as that:

$$\rho[H(\mathbf{w}^*, \mathbf{x}), h(\mathbf{x})] \leq \rho[H(\mathbf{w}, \mathbf{x}), h(\mathbf{x})] \quad \text{.....(3.23)}$$

Where $\rho[H(\mathbf{w}, \mathbf{x}), h(\mathbf{x})]$, or distance function, is a measure of approximation quality between $H(\mathbf{w}, \mathbf{x})$ and $h(\mathbf{x})$. When the fit is judged according to the sum of squared difference taken for the set for trained examples $\{\mathbf{x}\}$, the distance has a form of sum of squared errors.

3.1.5.2 *Supervised and Unsupervised Learning*: -

In supervised learning it was assumed that at each instant of time when the input is applied, the desired response $\mathbf{d}$ of the system is provided by the teacher. This is illustrated in Figure (3.21a). The distance $\rho[\mathbf{d}, \mathbf{o}]$ between the actual and the desired response serves as an error measure and is used to correct network parameters externally. Since adjustable weights were assumed, the teacher might implement a reward-and-punishment scheme to adapt the network's weight matrix $\mathbf{W}$. For instance, in learning classifications of input patterns or situations with known responses. This mode of learning is very pervasive. Also, it used in many situations of natural learning. A set of input and output patterns called a training set is required for this learning mode.

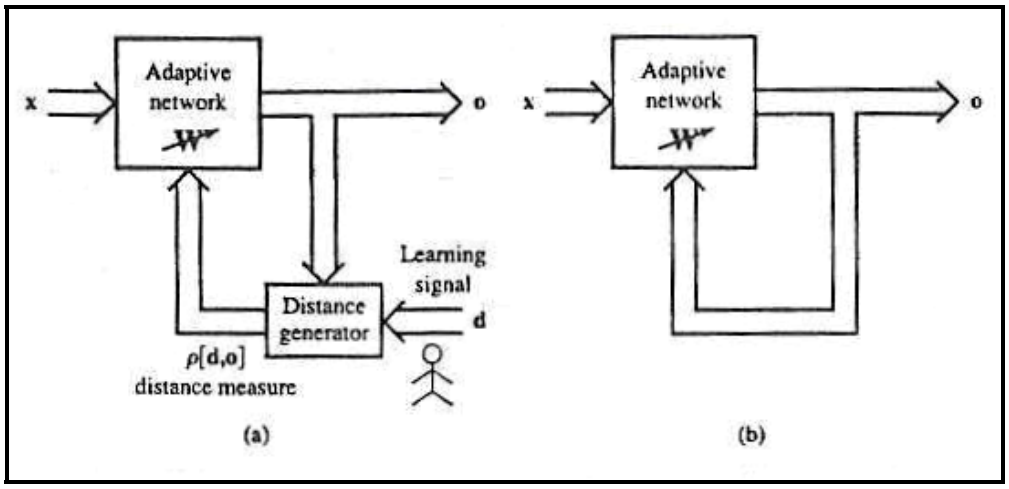


Figure (3.21): Block diagram for explanation of basic learning models: (a) supervised learning and (b) unsupervised leaning.

Typically, supervised learning rewards accurate classifications or associations and punishes those, which yield instance responses. The teacher estimates the negative errors gradient direction and reduces the error accordingly. In many situations, the inputs, outputs and the computed gradient are deterministic, however, the minimization of error proceeds over all its random realizations. As a result, most supervised learning algorithms reduce to stochastic minimization of error in multi-dimensional weight space.

Figure (3.21b) shows the block diagram of unsupervised learning. In learning without supervision, the desired response is not known; thus, explicit error information can't be used to improve network behavior. Since no information is available as to correctness or incorrectness of responses, learning must somehow be accomplished based on observations of responses to inputs that have marginal or no knowledge about. For example, unsupervised learning can easily result in finding the boundary between classes of input patterns distributed as shown in Figure (3.22). In a favorable case, as in Figure (3.22a), cluster boundaries can be found based on the large and representative sample inputs. Suitable weight self-adaptation mechanisms have to be embedded in the trained network, because no external instructions regarding potential clusters are available. One possible network adaptation rule is: A pattern added to the cluster has to be closer to the center of the cluster than the center of any other cluster.

Unsupervised learning algorithm use patterns, which are typically redundant raw data having no labels regarding their class membership, or associations. In this mode of learning, the network must discover for itself any possible existing patterns, regularities, separating properties, etc. while discovering these; the network undergoes change of its parameters, which is called self-organization.

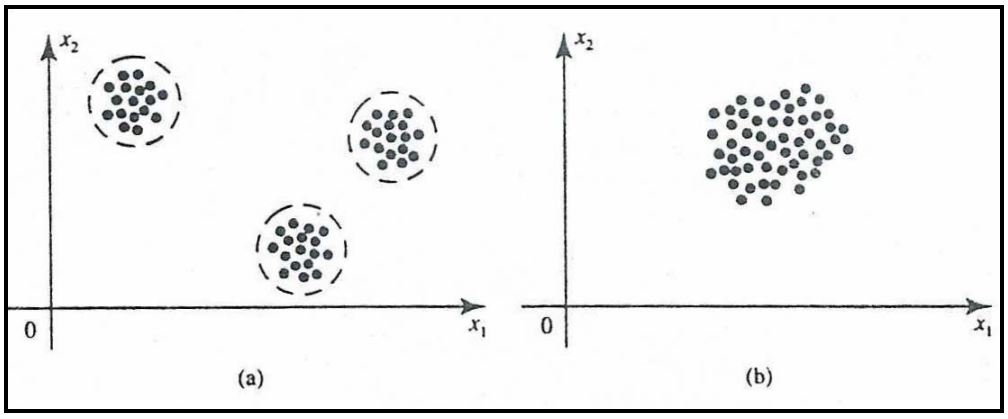


Figure (3.22): Two-dimensional patterns: (a) clustering and (b) no apparent clustering.

The technique of unsupervised learning is often used to perform clustering as the unsupervised classification of objects without providing information about the actual learning classes. Finally, learning is often not possible in an unsupervised environment, as would probably be true in the case illustrated in Figure (3.22b) showing pattern classes not easily discernible even for a human.

3.1.6 Advantages and Disadvantages of ANNs:-

3.1.6.1 *Advantages of ANNs*: -

Artificial Neural Networks offers valuable characteristics unavailable together elsewhere, some of these advantages are:

First, they infer solutions from data without prior knowledge of the regularities in the data; they extract the regularities empirically. **Second**, these networks learn the similarities among the patterns directly from instances or examples of them. ANNs can

modify their behavior in response to the environment (i.e. shown a set of inputs with corresponding desired output, they self-adjust to produce consistent responses). **Third**, ANNs can generalize from previous examples to new ones. Generalization is useful because real-world data are noisy, distorted, and often incomplete. **Fourth**, ANNs are also very good at the abstraction of essential characteristics from input containing irrelevant data. **Fifth**, they are non-linear, that is, they can solve some complex problems more accurately than linear techniques do. **Finally**, ANNs are highly parallel. They contain many identical, independent operations that can be executed simultaneously, often making them faster than alternative methods.

3.1.6.2 *Disadvantages of ANNs*: -

ANNs also have several drawbacks for some applications. **First**, they may fail to produce a satisfactory solution, perhaps because there is no learnable function or because the data set is insufficient in size. **Secondly**, the optimum network geometry as well as the optimum internal network parameters are problem dependant and generally have to be found using a trail-error process. **Finally**, ANNs cannot cope with major changes in the system they are trained (calibrated) on a historical data set and it is assumed that the relationship learned will be applicable in the future. If there were any major changes in the system, the neural network would have to be adjusted to the new process.

3.1.7 Short Reviewing of Use of ANNs in the Field of Forecasting of Water Resources: -

In recent years, ANNs have become extremely popular for prediction and forecasting in a number of areas, including finance, power generation, medicine, water resources, and environmental science. Although the concept of artificial neurons was first introduced in 1943 by McCulloch and Pitts (Zurada, 1995), research into applications of ANNs has blossomed since the introduction of the backpropagation training algorithm for feedforward ANNs in 1986 by Rumelhart (Zurada, 1995). ANNs was thus being considered a fairly new tool in the field of prediction and forecasting. In the following

section, a review of 54 papers (43 papers were already reviewed and reported by Maier,1999, and the rest 11 papers were reviewed by the author, denoted by (*) in the bibliography) dealing with the use of neural network models for prediction and forecasting of water resources variables is undertaken in terms of modeling process adopted.

The modeling steps investigated including selection of performance criteria, data division & preprocessing, determination of model inputs, selection of network architecture, optimization of connection weights and validation. The major features of the models investigated were summarized below, including background information (variable modeled, location of application, model time step, forecast length), information about the data used (data type(i.e. real or synthetic) normalization rang, number of training samples, number of testing samples) information about network architecture(connection type, method used to obtain optimal network geometry, number of nodes per layer), information about the optimization algorithm used (optimization method, internal network parameters, (method used to optimize internal network parameters, hidden layer transfer function, learning rate, momentum value, epoch size, initial weight distribution range)) and the stopping criterion adopted. Not all of the investigated features were given explicitly in the papers and some have been inferred from the information provided.

The papers were grouped according to variable modeled and in ascending order of publication year within each of these groups. A distribution of papers by year of publication is shown in Table (3.1) it can be seen that the first papers were published in 1992 and that there has been a marked increase in the number of publications in the last few years. Forecasting of flow (including rainfall runoff, streamflow and reservoir inflows) was the focus of 26 papers while rainfall forecasting was the subject of a further 16 papers. In the one of the latter papers dated 1997 by Loke et al, runoff coefficients were estimated in addition of rainfall and in one of the former papers by Clair and Ehrman in 1996, carbon and nitrogen levels, as well as streamflow, were predicted. There are nine papers that deal with the forecasting of water quality variables such as algal

concentration (3), cyanobacterial concentration (3) and salinity level (3). In the one of the later papers by Bastarache et al. in 1997, PH was predicted in addition to salinity. The remaining three papers are concerned with the prediction of water levels (including water table level and river stage).

The modeling time step used the range from one minute to one year. Prediction (i.e. forecast length of 0 and forecasting (forecast length >0) were the focus of an approximately equal number of papers.

Real data were used in 46 papers, while data were generated synthetically in the remaining eight papers. It should be noted that in two papers belonging to the later group by Karunanithi et al. in 1994; and Tsintikidis et al. in 1997(Maier et al, 1999), real data were used in some of the validation sets. Scaling of the data to a range that is commensurate with the transfer function in the output layer was only discussed in 20 of the 54 papers.

Table (3.1): Distribution of papers reviewed by publication date

Year of publication	No. of Papers
2000	2
1999	3
1998	13
1997	19
1996	7
1995	4
1994	2
1993	2
1992	2

Feedforward network were used in 47 papers. The exceptions are by Chow and Cho in 1997; (Hsu et al. 1997), in which recurrent and hybrid architectures were used, respectively. Network connectivity was not discussed by many authors, although it is probably safe to assume that the majority of networks were fully connected. Networks with one hidden layer were used in the majority of case studies, although networks with two and three hidden layers were also used. Inmost instances, no reason for the number of hidden layers used was given.

The standard backpropagation algorithm was used to optimize the connection weights in 42 of the 54 papers reviewed. (Hsu et al. 1995) used the LLSSIM algorithm, which combines a linear least squares procedure with the multistart simplex algorithm, whereas the modified counter propagation network in (Hsu et al. 1997) used an unsupervised learning rule to train a Self Organized Feature Map (SOFM) layer and a supervised method based on linear least-squares estimation to optimize the weights in a modified Grossberg linear layer. Shamseldin in 1997 used a conjugate gradient method, to speed up the training process.

3.2 Other (Traditional) Numerical Simulators:-

In order to give an idea about how the ANNs model performance during this study, it was suggested to use other numerical simulators to compare their results with those found by using ANNs. The suggested models were; the Total Linear Model (TLM), the Linear Perturbation Model (LPM), and the Autoregressive Model (AR). Although these models are considered as a traditional modeling approach, but all of it including ANNs were classified as system-theoretic approach "black box" (Hsu et al., 1993), since they directly map between the input and the output without detailed consideration of the internal structure of the physical process.

3.2.1 The Total Linear Model (TLM):-

This model assumes a linear relationship between the total input (rainfall or inflows or both) series and the corresponding total outflow series.

For the time series of input and output data, with length n and time interval delta t, the linear multiple-input single-output relationship may be expressed, in terms of the pulse responses, as a series of linear algebraic equations see Equation (3.24). The basic theory is due to Liang and Nash in 1988 (Abdo, 1992).

$$y_t = \sum_{j=1}^{J} \sum_{k=1}^{m(j)} h_k^{(j)} x_{t-k+1}^{(j)} + u_t \quad \text{.....(3.24)}$$

where $x^{(j)}$ is the j^{th} input time series; y is the output time series; and $h^{(j)}$ and $m(j)$ are the unit pulse response ordinates and memory length of he system corresponding to the j^{th} input series $x^{(j)}$, respectively and u_t is a disturbance term.

A least square solution for the $h(j)$ may be obtained as follows:

Rewriting Equation (3.24) in terms of matrix form as:

$$\mathbf{Y} = \mathbf{X}^{(1)}\mathbf{H}^{(1)} + \mathbf{X}^{(2)}\mathbf{H}^{(2)} + \ldots + \mathbf{X}^{(J)}\mathbf{H}^{(J)} + \mathbf{U} \quad \text{.....(3.25)}$$

where $\mathbf{Y}$ is an $(n,1)$ column vector of the output series; $\mathbf{X}^{(j)}$ is an $[(n.,m(j)]$ matrix of the j^{th} input series, and $\mathbf{H}^{(j)}$ is an $[m(j),1]$ column vector of the pulse response ordinates corresponding to the j^{th} input series. Equation (3.25) may be also written as:

$$\mathbf{Y} = \mathbf{XH} + \mathbf{U} \quad \text{.....(3.26)}$$

where $\mathbf{X}$ is an (n,M) matrix of the form

$$\mathbf{X} = [\mathbf{X}^{(1)}\mathbf{X}^{(2)} \ldots \mathbf{X}^{(J)}] \quad \text{.....(3.27)}$$

and $\mathbf{H}$ is an $(M,1)$ column vector of length

$$M = \sum_{j=1}^{J} m(j) \quad \text{.....(3.28)}$$

which is obtained by writing each $\mathbf{H}^{(j)}$ vector as an element of a column vector $\mathbf{H}$, so that it has the form

$$\mathbf{H} = [\mathbf{H}^{(1)}\mathbf{H}^{(2)} \cdots \mathbf{H}^{(J)}] \quad \text{.....(3.29)}$$

the ordinary least square estimate of H is given by

$$\hat{\mathbf{H}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \quad \text{.....(3.30)}$$

where $\mathbf{X}^T$ is the transpose of matrix $\mathbf{X}$.

3.2.2 The Linear Perturbation Model (LPM):-

The LPM concept was developed and explored by Nash and Barsi in 1983 in the context of rainfall-runoff modeling, assuming that *in any one year in which the rainfall exactly follows the seasonal expectation, the discharge hydrograph will similarly follow its expectation, and in other years departures from the seasonal expected values in rainfall and discharge will be linearly related.* The model is therefore a marriage between two well established concepts, one in time series analysis (i.e. seasonal component identification) and the other in deterministic systems analysis (i.e. classical unit hydrograph identification) (Abdo et al,1992).

The seasonal means of output (and similarly for the input) on date d are obtained in the calibration period by:

$$q_d = \frac{1}{n}\sum_{r=1}^{n} q_{d,r} \quad \text{.....(3.31)}$$

where n is the number of years in the calibration period and $q_{d,r}$ is the output on date d for the r^{th} year of calibration data. In practice, the q_d so obtained may be smoothed by Fourier Smoothing and the resulting seasonal expectation of output (as input) is the simplest model, and is called the Seasonal Model (SM) developed by Garrick 1978 (Eltahir, 1987), (Abdo et al, 1992).

The formula for smoothing the discharge seasonal means is give below

$$q_d = \bar{q} + \sum_{j=1}^{182} \{A_j \cos(\frac{2\pi jd}{365}) + B_j \sin(\frac{s\pi jd}{365})\} \quad \text{, For d=1, 2… 365} \ldots\ldots (3.32)$$

where

$\bar{q} \equiv$ The mean discharge,

$A_j, B_j \equiv$ Are the Fourier coefficients, and j is the serial number of the harmonic.

$$\bar{q} = \frac{1}{365} \sum_{d=1}^{365} q_d \quad \ldots\ldots (3.33)$$

$$A_j = \frac{2}{365} \sum_{d=1}^{365} q_d \cos(\frac{2\pi jd}{365}) \ldots\ldots (3.34)$$

$$B_j = \frac{2}{365} \sum_{d=1}^{365} q_d \sin(\frac{2\pi jd}{365}) \ldots\ldots (3.35)$$

In order to smooth the curve formed by joining the 365 daily seasonal means, the Fourier series defined by Equation (3.32) must be truncated, i.e. only few out of the 182 harmonics must be used. The most powerful harmonic was selected based on visual inspection of the actual seasonal means and the smoothed seasonal one time series plots.

The LPM concept can be explained through the following equations:

$$y_t = q_t - q_d \ldots\ldots (3.36)$$

$$x_t = i_t - i_d \ldots\ldots (3.37)$$

where

$y_t \equiv$ The output series of departure at day t,

$x_t \equiv$ The input series of departure at day t.

From the TLM concept, estimate the output series of departure y_t from the input series of departure x_t, which can be explained by Equation (3.24). And the final estimated output can be obtained as follows:

$$q_t = q_d + \hat{y}_t \quad \ldots\ldots (3.38)$$

where

$q_t \equiv$ Estimated output,

$q_d \equiv$ Smoothed seasonal means of the output series,

$y_t \equiv$ Estimated output series of departures.

3.2.3 The Modified Linear Perturbation Model (LPM):-

As described in the above section, the LPM simulates linearly the departures from the seasonal expectation for the input and output series. The modification carried during this study, is concerned with the use of a different model to simulate the output and input departures series. Two models were suggested namely; the Autoregressive Model (AR) which is described in the following section, and a combination between the Autoregressive Model and the Total Linear Model (AR-TLM). The second suggested model can be expressed by the following equation:

$$y_t = \sum_{j=1}^{J} \sum_{k=1}^{m(j)} h_k^{(j)} x_{t-k+1}^{(j)} + \sum_{k=1}^{m} \phi_k y_{t-k} + u_t \quad \text{..... (3.39)}$$

where $x^{(j)}$ is the j^{th} input time series; y_t is the output time series at time t;

and $h^{(j)}$ and $m(j)$ are the unit pulse response ordinates and memory length of the system corresponding to the j^{th} input series $x^{(j)}$, respectively; and $\phi_{(k)}$ is the autoregressive parameter at $t-k$ time step and m is the memory length of the system corresponding to the output series, and u_t is a disturbance term.

3.2.4 The Autoregressive Model (AR):-

In a linear autoregressive model, the value of a random variable under consideration at one time period is equated (correlated) to the weighted sum of a pre-assigned number of the random variable at earlier time periods. The method dates back to the classical work of Markov in 1907. The basic theory is due to Yule in 1921 & 1927 and Wolker in 1931 (Kottegoda, 1980).

The p^{th}-order linear autoregressive model $\mathrm{AR}(p)$ of a random variable X at one time period with its value p time period earlier can be shown by the following equations,

$$X_t = x_t - \bar{x} \quad \text{.....(3.40)}$$

where

$\bar{x} \equiv$ The mean of the series,

$X_t \equiv$ Departures of the i^{th} value from the mean of the series.

And

$$X_t = \sum_{i=1}^{p} \phi_{p,i} X_{t-i} + \varepsilon_t \quad \text{.....(3.41)}$$

where

$\phi_{p,i}\ for\ \ i = 1,2...p \equiv$ The autoregressive parameters or weights.

$\varepsilon_t \equiv$ Random component, white noise and usually assumed to be normally distributed with mean zero and variance σ_ε^2.

The variance of the white noise component is given by

$$\sigma_\varepsilon^2 = \sigma_x^2 .(1 - R^2) \quad \text{.....(3.42)}$$

where

$R^2 \equiv$ The multiple coefficient of determination between X_{i+1} and $X_i, X_{i-1}, \cdots, X_{i-p+1}$.

And

$$\varepsilon_t = \sigma_x t_{t+1} \sqrt{1 - R^2} \quad \text{.....(3.43)}$$

The order of the autoregressive model can be determined by using the Partial Autocorrelation Function (PAF) (Kottegoda, 1980). The set of parameters $\phi_{1,1}, \phi_{2,2}, ...\phi_{i,i}$

which are the last coefficient of the autoregressive models of order 1, 2,…. i, respectively constitute the (PAF). In general, the partial autocorrelation $\phi_{p,p}$ is the autocorrelation remaining in the series after fitting a model of order p-1 and removing the linear dependences. It is important to note that, whereas the autocorrelation function decreases to zero only at an infinite lag, in an autoregressive process of order p the partial autocorrelations $\phi_{k,k}$, are theoretically zero for $k > p$. Therefore, the sample PAF is an important tool in determining the order of the model if the serial correlation function suggests that the process could be approximated by a linear autoregressive model.

C h a p t e r 4

MODEL DATA DEVELOPMENT

The data used during this study consist of the following: watershed data, meteorological data, and hydrological data.

4.1 Watershed Data: -

4.1.1 Description of the Study Area: -

The Atbara River is the most northern and the last tributary of the Nile, is a strongly seasonal river, which enters the Main Nile at about 320km downstream of Khartoum. Its headwaters originate in the northwestern Ethiopian Highlands, 880km long. The entire Atbara sub-basin is quite large. It counts 6,675 METEOSAT pixels, which corresponds to 166,875 km² (about 8º-15ºN, 34º-39ºE) (Roskar, 2000). The greater part of its catchment is suited in the Ethiopian and Eritrea. The highest points in the catchment reach more than 3,500 m, whereas the eastern watershed of Atbara is, for the most part, more than 2,500 m high. The River relies totally on many small tributaries, of which Takazze or the Setit is the principal one. Above the Setit junction, the Atbara receives a number of tributaries of which Bahr el Salam is the principal (Karyabwite, 2000) see Figure (4.1).

4.1.2 Watershed Discritization and River Network:-

The data used to identify the river network and watershed discretization during this study was reported by Josef Roskar, in October 17th 2000 former chief technical advisor of the Monitoring, Forecasting, and Simulation of the Nile-Egypt project (MFS) and by the Flood early Warning System (FEWS)-Ministry of Irrigation -Sudan.

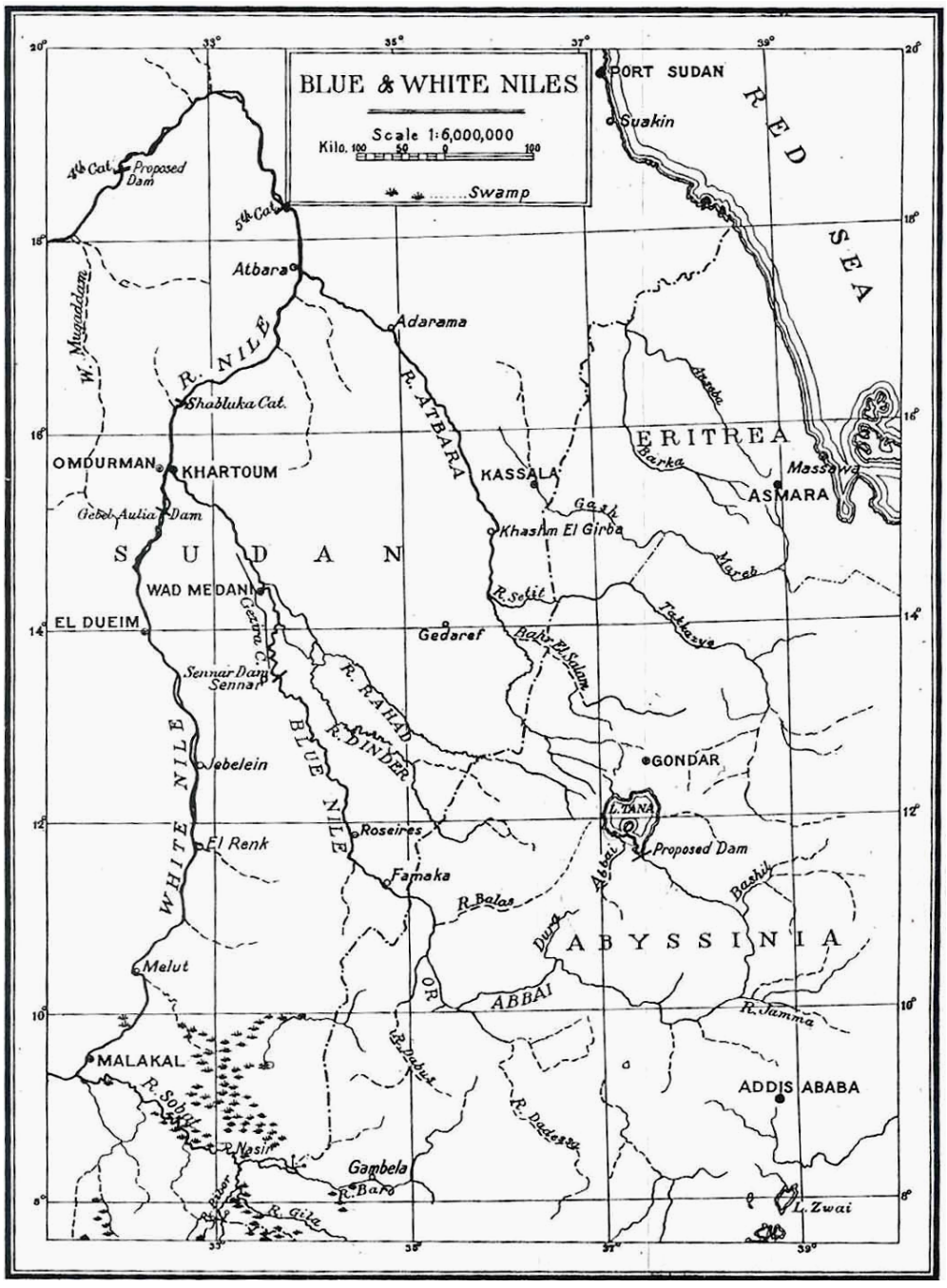


Figure (4.1): The river drainage network for the study area (Nile Water Department, 1996).

The data stored in the Nile basin Hydrometeorological Information System. The basic resolution of the girded data is a METEOSAT pixel (5km*5km in the sub-satellite point). Eight major sub-basins within the Nile Basin were identified and selected on the basis of

watershed drainage divides, sub-basin characteristics, and the location of river gauging sites. The data studied to identify the sub-basins was for the 1940-1995-time period. The main characteristics of the sub-basins were shown in Table (4.1).

The two major sub-basins Blue Nile and Atbara River were divided into minor sub-basin and delineated in Figure (4.2) (Water Resources Department, 1992).

Table (4.1): The main hydrological characteristics of the Nile river major sub-basins (Roskar, 2000).

Sub-basin	Outlet	No. of Pixels	Area (km^2)	Average rainfall (mm/year)	Total rainfall (km^3/year)	Average runoff (km^3/year)	Runoff/ Rainfall (%)
Lake Victoria	Jinja	9,546	238,650	1,295	309.05	30.97	10.02
Equatorial Lakes	Mongalla	7,784	194,600	1,198	233.13	6.54	2.81
Sudd Area	Malakal	5,577	139,425	923	128.69	Loss of flow	
Bahr al-Ghazal	Lake No	13,215	330,375	970	320.46	0.50	0.16
Sobat River	Helit Dolieb	7,451	186,275	1,057	196.89	13.66	6.94
Ethiopian Highlands	Diem	5,676	141,900	1,346	191.00	47.44	24.84
Blue Nile in Sudan	Khartoum	4,847	121,175	573	69.43	2.00	2.88
Central Sudan	Khartoum			Semi-arid		Loss of 4.5	
Atbara River	Atbara Kilo3	6,675	166,875	553	92.28	10.93	11.84
Entire Nile Catchment	Dongola	61,103	1,527,500	1,010	1542.78	84.71	5.49

4.2 Meteorological Data: -

4.2.1 Rainfall Data:-

Six years of lumped daily rainfall records were used during this study. FAO-Focal Point Institute-Water Resources Project-Ministry of Irrigation-Sudan supplies the data set. It includes daily rainfall records for ten stations (two different data sets (a) and (b)

from different sources with different length of records). Table (4.2) shows the station information (WMO Classification), Figure (4.2) shows where these stations were located, and Tables (4.3) & (4.4) shows general description of these records.

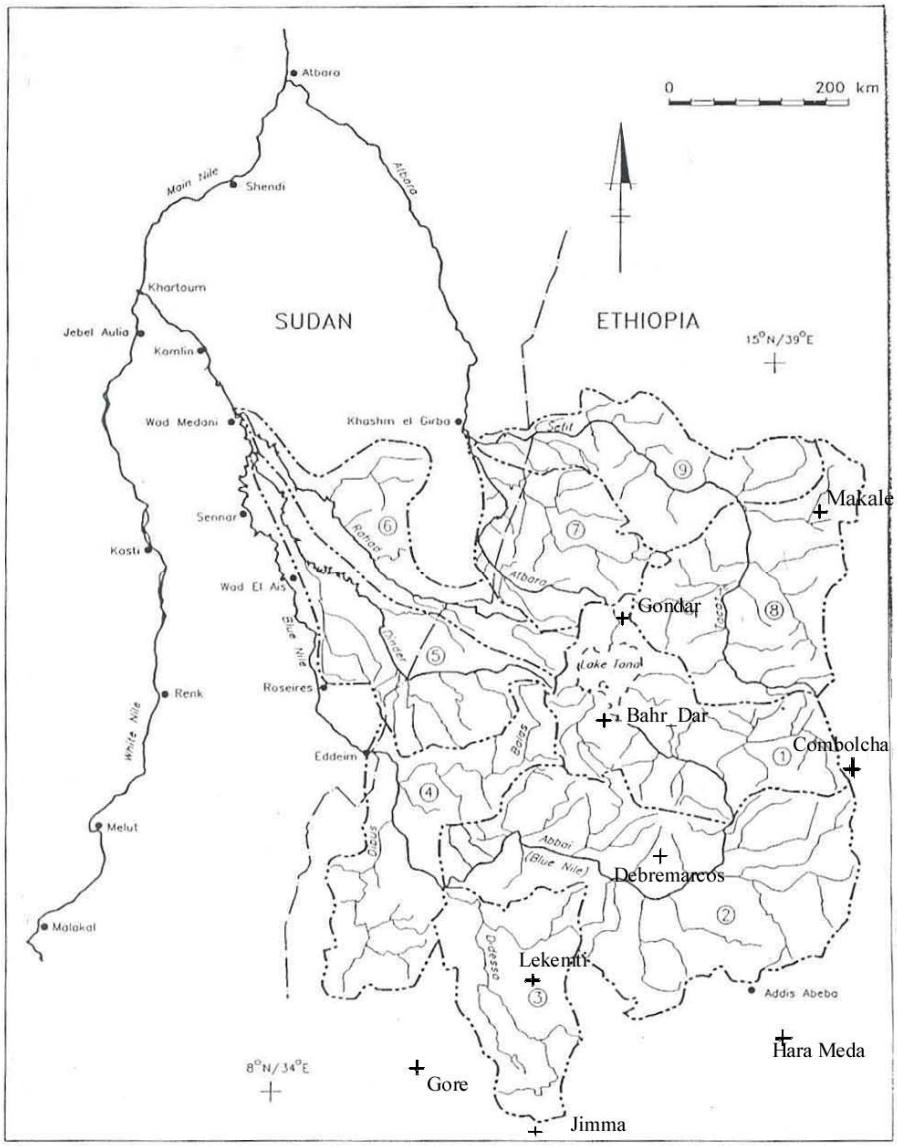


Figure (4.2): Watershed discretization and rainfall station location (Nile Water Department, 1992).

Table (4.2): Meteorological Station Information Lookup.

WMO Index Number	ICAO Location Indicator	Station Name	Country	WMO Region	Station Position (dms)	Station Elevation(H_a) (m)	Upper Air Elevation(H_p) (m)
63330*	HAMK	Makale	Ethiopia	1	13-30N 39-29E	2212	2212
63331*	HAGN	Gondar	Ethiopia	1	12-32N 37-26E	1985	1966
63332*	HABD	Bahar Dar	Ethiopia	1	11-36N 37-24E	1820	1805
63333*	HADC	Combolcha	Ethiopia	1	11-05N 39-43E	1864	1916
63334*	HADM	Debremarcos	Ethiopia	1	10-21N 37-43E	2476	2440
63340	-	Lekemti	Ethiopia	1	09-03N 36-36E	1950	1950
63402*	HAJM	Jimma	Ethiopia	1	07-40N 36-50E	1676	1577
63403*	HAGR	Gore	Ethiopia	1	08-10N 35-33E	1974	1974
63450*	HAAB	Addis Ababa	Ethiopia	1	08-59N 38-48E	2355	2324
63451*	HAHM	Harar Meda	Ethiopia	1	08-44N 39-00E	1876	1876

*= **This station participates in the WMO Regional Basic Synoptic Network.**

Table (4.3): General description of the rainfall records that extend for the years (1990 to 1998) (data set a).

Station	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
GONDAR*	2519	76.74%	768	23.36%	3287	100.0%
MAKALE*	2079	63.25%	1208	36.75%	3287	100.0%
BAHAR DAR*	2777	84.49%	510	15.51%	3287	100.0%
COMBOLCHA	2613	79.49%	674	20.51%	3287	100.0%
DEBRE MARKOS	2510	76.36%	777	23.64%	3287	100.0%
ADDIIS ABABA	2946	89.63%	341	10.37%	3287	100.0%
LEKEMETI	2017	61.36%	1270	38.64%	3287	100.0%
GORE	1544	46.97%	1743	53.03%	3287	100.0%
JIMMA	2072	63.04%	1215	36.96%	3287	100.0%
HARRAR MEDA	2178	66.27%	1109	33.73%	3287	100.0%

***= shows the station considered in the study**

Table (4.4): General description of the rainfall records which extend for the years (1993 to 1998) (data set b).

Station	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
GONDAR*	2188	99.9%	3	0.14%	2191	100.0%
MAKALE*	2191	100.0%	0	0.0%	2191	100.0%
BAHAR_DAR*	2187	99.8%	4	0.18%	2191	100.0%
COMBOLCHA	2191	100.0%	0	0.0%	2191	100.0%
DEBRE_MARKOS	2147	98.0%	44	2.01%	2191	100.0%
ADDIS_ABABA	2071	93.4%	120	5.48%	2191	100.0%
JIMMA	2190	99.95%	1	0.05%	2191	100.0%

***= shows the station considered in the study**

Appendix (A4.1) shows graphical plots and some statistical measures of data for the stations considered on the study.

4.2.1.1 *Quality of Rainfall Data*:-

In order to have an idea about the quality of data used in the current study; the adequacy of the number of rain gauges available in the region considered and the homogeneity of the records were discussed below.

4.2.1.1.1 *Adequacy of the number of rain gauges in the region considered*:-

Three rainfall stations were used in this study. The area covered was about 166,875 km² with station density of one station per 55,625 km².

To give an idea about how poor the rain gauge network at the considered region is, the National Center for Atmospheric Research (NCAR) gives in Table (4.5) the station densities for other parts of the world (Eltahir, 1987).

Also Table (4.6) (Rodda et al, 1976) gives recommendations for minimum numbers of rain-gauges.

Table (4.5): Rainfall Station Densities for some Hydrologicaly similar Areas (Eltahir, 1987).

Continent, country or region	Area (km^2)	No. of stations	Station density (km^2/station)
Australia	8,000,000	14,000	571
India	3,268,000	4,000	817
Africa	30,300,000	1,087	27,875
Sudan	2,560,000	44	58,182
Region of study*	166,875	3	55,625

***= shows the region considered in this study.**

Table (4.6): Recommended minimum densities for rain-gauge networks in reservoired basins (Rodda et al, 1976)

Area		Rain-gauge type		
($mile^2$)	(Km^2)	Daily	Monthly	Total
0.8	2	1	2	3
1.6	4	2	4	6
7.8	20	3	7	10
15.7	41	4	11	15
31.3	81	5	15	20
46.9	122	6	19	25
62.5	162	8	22	30

This fact was later observed in deciding how to use the records available, adopting a regional mean or treating each station record separately. It is thought that a reasonable compromise may be achieved by adopting a regional mean.

4.2.1.1.2 *Homogeneity of Rainfall Data*:-

The data used were subjected, like every other data, to errors in all phases of measuring, transmission, recording and processing. These errors can be classified into random, systematic and gross errors.

Random errors are always in data and are assumed to be symmetrically distributed around the true value; the standard deviation of this distribution is a measure of the magnitude of those errors.

Gross errors are those due to unseen mistakes, usually in the processing phase, like replacing one station record with a record from another station. Such mistakes may lead to serious damages for any analysis but, can easily be detected.

Systematic errors in rainfall data may be caused by any factor which may change precipitation catch of a rain gauge namely: changes in gauge location, exposure, instrumentation or observational procedure. A record with systematic errors is regarded as nonhomogeneous, or some times described as in consistence record.

The World Meteorological Organization (WMO) defines the homogeneity of a data set as "uniform representativeness (in time) of the data to conditions prevailing in a large geographical area", (Eltahir, 1987).

For the data used in this study the following checks were applied:

1- The daily rainfall series where checked for observing the seasonal pattern characteristics of rainfall at the region. Chronological charts were developed for daily rainfall series for each station and checked visually for outliers.

2- The overlapping period between the two data sets gave a chance to check how consistent the two data sets are. The two sets were found to agree in general.

3- Double-Mass analysis was carried for all the records and the results obtained in favor of, at least, relative homogeneity of the records.

Appendix (A4.2) shows the results of this analysis. The analysis procedure is elaborated in the next subsection.

4.2.1.1.2.1 *Double-Mass Analysis***:-**

Homogeneity of any rainfall station record compared to average rainfall amounts at the whole region can be satisfied if the daily rainfall amounts at that station are showing a uniform linear relation to the regional average amounts. This can be expressed mathematically by the following equation,

$$R_x(I) = C * R_{av}(I) + e(I) \quad \text{.....} (4.1)$$

where

$R_x(I) \equiv$ Daily rainfall amounts at station x for day I ,

$R_{av}(I) \equiv$ Daily rainfall amounts averaged for the whole region for day I ,

$e(I) \equiv$ Amount of white noise in the station daily rainfall for day I .

A possible way of testing the relative homogeneity of any station record is to plot R_x vs. R_{av} and check how well the above equation is satisfied, but since both R_x and R_{av} do fluctuate around a constant mean, their graphical display may not be the appropriate test for homogeneity. Instead the plot of the accumulated daily amounts at the station against accumulated daily amounts averaged over the whole region has got a better graphical display and is usually used for testing homogeneity. This is known as the double-mass analysis.

The theoretical relation can be easily obtained, for any record starting at year M and for any following year N , adding up Equation (4.1), with I ranging from M to N get,

$$\sum_{I=M}^{N} R_x(I) = C * \sum_{I=M}^{N} R_{av}(I) + \sum_{I=M}^{N} e(I) \quad \text{.....(4.2)}$$

If in any following year M', for any reason, the gauge catch is changed then the station record will no longer be homogeneous and Equation (4.1) will be changed into,

$$R_x(I) = C' * R_{av}(I) + e(I) \quad \text{.....(4.3)}$$

Following which equation (4.2) will be changed into,

$$\sum_{I=M}^{N} R_x(I) = C'' + C' * \sum_{I=M}^{N} R_{av}(I) + \sum_{I=M}^{N} e(I) \quad \text{.....(4.4)}$$

This will appear in the plot as a sudden change in the slope of the line representing the above relation from C to C'. C" is given by the following equation,

$$C'' = (C - C') * \sum_{I=M}^{N} R_{av}(I) \quad \text{.....(4.5)}$$

For the three stations, which were used in the study, the double-mass analysis was carried for each station assuming the rest of the stations (including other seven station nearby the study area), although no station was guaranteed to be homogeneous. Since all tests gave results in favor of reasonable homogeneity of each station record with respect to the rest of the stations, the three stations were accepted to be homogeneous.

The relation between the accumulated daily amounts at any station and accumulated daily amounts averaged for the whole region was also investigated by linear regression of the form similar to the theoretical assumed relation. The proportion of the initial variance explained by the regression equation, which is the coefficient of determination, provides a numerical measure of the homogeneity of the station record. Table (4.7) shows those percentages of the explained variances for the stations used in the study.

4.3 Hydrological Data: -

4.3.1 River Stage Data:-

Flood Early Warning System Unit - Data Center - Ministry of Irrigation & Water Resources – Sudan supplied observed daily river stages for three stations, which represent the variation in the stages for the Atbara River and its tributaries. Wed el Heliew, Kubur, and Atbara Kilo_3 are the names of the stations located at Setit, Atbara, and Atbara River respectively (Figure (4.3)). The records extend for the years 1993 to 1998. Table (4.8) shows a general description of these records. In addition, there are also scattered records for the discharges with the corresponding stages from which the rating curves for these stations could be generated. Table (4.9) shows a general description of these records.

Table (4.7): Results of the linear regression of the accumulated daily rainfall amounts at each station and accumulated daily regional average

Station	% of the variance explained by the linear relation	Ratio(C) of station annual rainfall to the regional average rainfall
GONDAR*	99.47	0.956
MAKALE*	99.3	0.509
BAHAR_DAR*	99.75	1.119
COMBOLCHA	99.66	0.984
DEBRE_MARKOS	99.87	1.200
ADDIS_ABABA	99.65	0.902
JIMMA	99.44	1.397

Table (4.8): General description of the river stage records for the stations considered during the study.

Station	Length of Record (days)					
Year	1993	1994	1995	1996	1997	1998

Wed El-Heleiw	365	340	365	366	365	354
Kubur	291	212	212	271	224	238
Kilo_3	142	55	72	127	92	146

Table (4.9): General description of the data used to develop rating relationship.

Station	**Length of Record (days)**					
Year	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**
Wed El-Heleiw	95	53	52	61	55	-
Kubur	5	17	30	-	21	-
Kilo_3	12	14	3	20	9	14

Appendix (4.3) shows time series plot and some statistical properties of the stage records at the three stations.

4.3.2 Khashm El-Girba Reservoir Data:-

Khashm El-Girba reservoir was mainly constructed for agriculture and power generation purposes in 1964. The initial storage capacity of the reservoir was about 1.3 milliard m^3 in a reservoir of 80 km length and total surface area of about 125 km^2. The location of the reservoir site was shown in Figure (4.3). Its total average annual inflow is about 11.8 milliards m^3, average flood of 2000 m^3/sec. and maximum discharge capacity at level 473 m (amsl) is 8.7 m^3/sec. the normal impounded level of the dam is 473 m (amsl) where as its exceptional catastrophically level is 474.5 m. the minimum operational level is 462 m (amsl) (recently raised to 463.5 m).

Six years of daily records for the up-stream and down-stream levels, main canal discharge and the dam release to the stream were provided in addition to 10 days volume

for the reservoir evaporation and power generation in million m^3. Table (4.10) shows general description of these records.

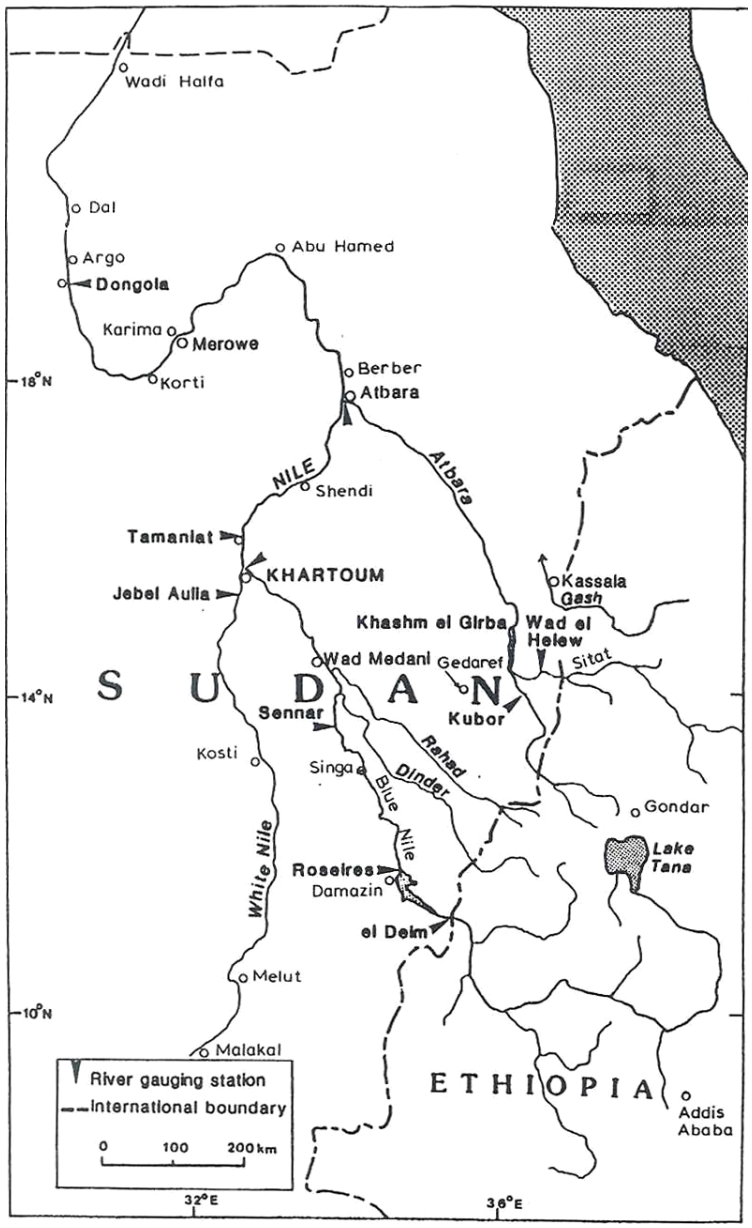


Figure (4.3): River gauging stations location map (Nile Water Department, 1992).

Table (4.10): General description of Khashm El-Girba reservoir records.

	Length of record					
	1993	**1994**	**1995**	**1996**	**1997**	**1998**
U/S Level	365	365	365	366	365	365
D/S Level	365	364	365	366	365	365
Release	231	146	253	249	178	292
Discharge into the Main Canal	362	359	243	366	365	365

Appendix (4.4) show some statistical measures of these records

4.3.3 Missing Stage Amounts: -

The filling of missing data points is a problem facing almost every research on real data. The data can be lost due to different reasons, problems in measuring, recording, transmission or processing can lead to recording the data point as missing value. All the data sets were checked for the missing value and it was about 30% of the data is missing due to problems in measurement or transmission. Worth mentioning that about 92% of the data reported as missing values were found at Kilo_3 station.

For filling in the missing stage amounts, information from other stations with complete data set and lay within the same hydrological region can be used (Varshney, 1979). Different models were suggested, since only the most efficient one should be used, an objective way for deciding this was thought to be necessary. The Nash and Sutcliffe efficiency criterion E_{ff}, (Equation (2.6)), which is used in comparing different models in

streamflow forecasting, was found to be a useful tool by which a decision can be made of which model to use in filling missing stage amounts.

4.3.3.1 *Models for Filling in Missing Stage Amounts*:-

Three forms of regression model were tried. They are all based on the concept of minimizing the summation of the squares of differences between the model estimates, using observed daily stage amounts at the two nearby stations, and the observed amounts at the station with the record to be filled, in order to obtain the parameters of the regression equation. These parameters may later be used for utilizing observations from the two stations in estimating a missing data points.

4.3.3.1.1 *Linear Relation with Constant*:-

The first form can be expressed by the equation

$$y_m(i) = k + a.x_a(i) + b.x_b(i) + e(i) \quad \text{.....(4.6)}$$

where

$y_m(i), x_a(i) and \quad x_b(i) \equiv$ The stage amounts at station m, a, b, respectively,

$k, a, \& b \equiv$ Constants,

$e(i) \equiv$ Random component at day i.

4.3.3.1.2 *Linear Relation without Constant*:-

The second form can be expressed by the following equation

$$y_m(i) = a.x_a(i) + b.x_b(i) + e(i) \quad \text{.....(4.7)}$$

where all variables and constants are as defined above.

4.3.3.1.3 *Linear Autoregression Relation without Constant*:-

The third form can be expressed by the equation

$$y_m(i) = a.y_{m-1}(i) + b.x_b(i) + e(i) \quad \text{.....(4.8)}$$

where all variables and constants are as defined above.

4.3.3.2 *Procedure for Filling in Missing Stage Amounts*:-

1- Replace missing values using a linear interpolation. The last valid value before the missing value and the first valid value after the missing value are used for the interpolation. If the first or last case in the series has a missing value, the missing value is not replaced.

2- Apply the model with the highest efficiency and stable parameters. For this study it was always the linear form of regression without constant.

3- Two stations lies within the same hydrological region of the station with missing values were tried and a stepwise approach was used to find the efficient model.

Appendix (4.5) shows, in detail, how the missing amounts at Kilo_3 station are filled in.

4.3.4 Estimation of Rating Curves: -

The data set supplied for river stages contains also rating relationships, which can be used in the transformation of stages to discharges. A thorough investigation and revision of these relations was carried to ensure its accuracy, since the quality of the data play a great role in producing efficient models and more reliable results.

This process was performed by developing new rating relationships using the same records that used to develop those provided. Many methods were suggested in the literature to be used in developing such relationships such as the logarithmic method, Steven's method and Manning formula (Varshney, 1979). The logarithmic method was selected to be used, since both the other two methods require additional information about the hydraulic section, which was not available. The Logarithmic method expresses the relationship between stage and discharge as follows:

$$q = b_1 (h - b_2)^{b3} \quad \text{.....(4.9)}$$

Where:

$q \equiv$ Discharge at stage h,

b_1 & $b_3 \equiv$ Constants for the station,

$b_2 \equiv$ Zero elevation of zero flow (approximated).

The optimum parameters values $b_1, b_2, \& b_3$ were determined by following a systematic trail and error procedure with the objective to minimize the sum of the squared differences between the estimated discharges $\hat{q}$ using the observed stages values, and the observed discharges at the station. For each year a rating relationship was developed for Wed El-Heleiw and Kubur station, but for Kilo_3 station one relation was developed for the whole period (1993 to 1998), since it had few data and it shows very poor quality.

The graphical and tabulated results shown in Appendix (A4.6), show that the newly developed rating relationships are more efficient and more reliable.

4.3.5 Khashm El-Girba Reservoir Water Balance Model:-

The inflows to Khashm El-Girba reservoir were the only available information downstream the junction of Setit and Atbara tributaries, that can be used to express the combined effect that was generated from the combination of the two hydrographs developed at stations Wed El-Heliew and Kubur located at the Setit and Atbara tributaries, respectively .

$$I - (R + M.C. + E + P) = \Delta S \quad \text{.....(4.10)}$$

$$\Delta S = S_i - S_{i-1} \quad \text{.....(4.11)}$$

where

$I \equiv$ The daily discharges into the reservoir (Inflows).

$R \equiv$ The daily discharges released from the dam into the main river stream.

$M.C. \equiv$ The daily discharges released from the dam into the main canal for irrigation purposes.

$E \equiv$ The daily evaporated volumes from the reservoir.

$P \equiv$ The daily discharges released from the dam for power generation purposes.

$\Delta S \equiv$ The daily changes in reservoir storage, which can be found by finding the differences between the storage at day i (S_i) and storage at day $i-1$ (S_{i-1}).

The calculated daily inflows from the model were evaluated; it was found that it contains illogical values because of the existence of negative discharges. These values may exist due to incorrect recording of the data. It were considered as missing values and replaced with linearly interpolated ones; Figure (4.4) shows the results of the calculated inflows hydrograph.

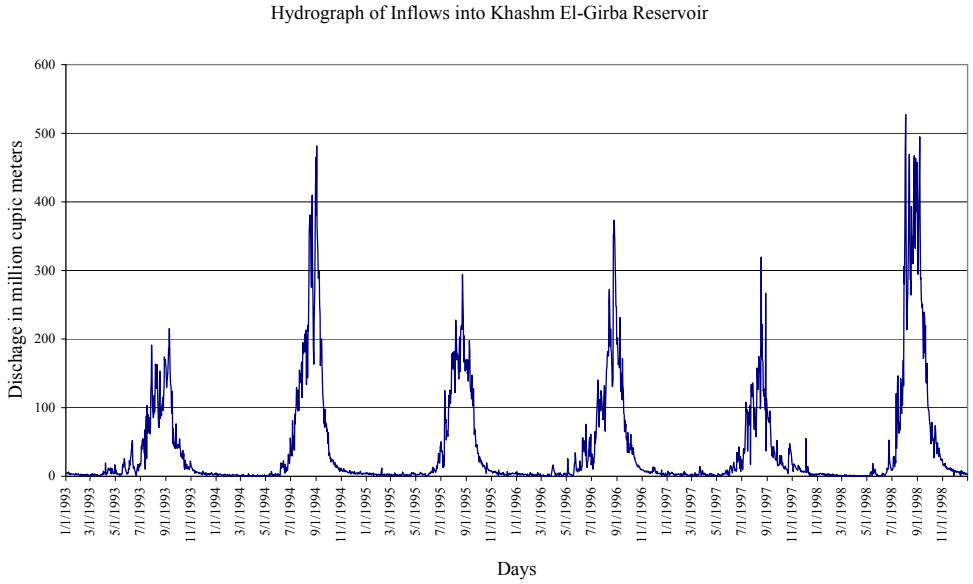


Figure (4.4): The hydrograph of the calculated inflows into Khashm El-Girba reservoir.

Chapter 5

CALCULATION, RESULTS AND DISCUSSION

5.1 General:-

This section elaborates how the forecasting models, detailed in chapter (3), were built and developed to simulate both rainfall-runoff process as well as channel routing process in the region being studied. The system being modeled is shown in Figures (4.2) and (4.3). It show that, the runoff measured at four stations namely; Wed El-Heleiw, Kubur, Khashm El-Girba dam and Kilo_3, were generated as a consequence of the rainfall events over the Ethiopian Highlands. Special consideration during the model building process was given to ANNs model. This is because, the study objectively concerns with the application of this recently developed modeling tool to be used in this field, and many issues such as those related to network architecture are still not well understood (Hsu et al., 1997), but some guidance can be obtained from the findings of other researchers.

5.2 ANNs Model Description:-

As in any prediction/forecasting model, the selection of appropriate model inputs is extremely important step to be followed in the model building process. However, in most ANN applications, little attention is given to this task, that is because ANNs belong to the class of data driven approaches, whereas conventional statistical methods are model driven. In the later, the structure of the model has to be determined first, which is done with the aid of empirical or analytical approaches, before the unknown model parameters can be estimated. Data driven approaches, on the other hand, have the ability to determine which model inputs are critical, so there is no need to "a priori rationalization about relationships between variables". However presenting a large number of inputs to ANN models, and relying on the network to determine the critical model inputs usually increases the network size. This has a number of disadvantages, such as decreasing processing speed and increasing the amount of data required to estimate the connection

weight efficiently. The problem is exacerbated in time series applications, where appropriate lags have to be chosen for each of input variables.

Therefore, in the current study a simplified methodology was adopted to determine the appropriate number of inputs and their lags for a simple univariate ANN model. The methodology involves the use of some traditional models namely; the Total Linear Model (TLM), the Linear Perturbation Model (LPM) and the Autoregressive Model (AR) with a systematic trial and error procedure.

5.2.1 Determination of Network Architecture:-

Network architecture determines the number of connection weights (free parameters) and the way information flows through the network. Determination of appropriate network architecture is one of the most important, but also one of the most difficult, tasks in the model building process.

5.2.1.1 *Type of Connection and Degree of Connectivity*:-

Reviewing the ANNs literature theories, two main ANNs structures were identified namely; FeedForward Networks (FFN) and Recurrent Networks (RN). Theoretical and practical studies (based on review of about 54 papers, see section 3.1.7) have indicated that the FFN (i) have been found to perform well in comparison with recurrent networks. (ii) They have been used almost exclusively for the prediction and forecasting of water resources variables. (iii) Their processing speed "is among the fastest of all models developed and used recently. (iv)Recurrent networks "do not provide clear practical advantages over feedforward nets with limited windows. Thus, from the above, it is thought that satisfactory results may be obtained from the use of a feedforward nets with fully connected networks.

5.2.1.2 *Network Geometry*:-

Network geometry determines the number of connection weights and how these are arranged. This is generally done by fixing the number of hidden layers and choosing the

number of nodes in each of these. The number of nodes in the input layer is fixed by the number of model inputs, whereas the number of nodes in the output layer equals the number of model outputs. Traditionally, optimal network geometries have been found by trail and error procedure. More recently, a number of systematic approaches for determining optimal network geometry have been proposed, including pruning and constructive algorithms. Constructive algorithms approaches were selected to be used. In the current study in the constructive algorithms the smallest possible network (i.e. one without hidden layers) is used at the start of training. Hidden layer nodes and connections are then added one at a time in an attempt to improve model performance. The major issues that need to be addressed in constructive algorithms include the way additional nodes are connected to the existing network, how to determine optimal connection weights (new and existing) once a node has been added and when to stop the addition of hidden nodes. Software called Pythia was used to perform this process supported by the Evolutionary Optimizer algorithm see Figure (5.1).

5.2.2 Input and Output Data Processing:-

The original input output data sets were linearly transformed (Scaled or normalized) to the interval (0.05-0.95). The reason for that, if the values are scaled to the extreme limits of the transfer function, the size of the weight updates is extremely small and flat spots in training are likely to occur (Maier et al., 1999)

Let a and A to be the minimum and the maximum values of the input data series, respectively; then an actual value x_a is transformed to x_t, to the interval (0.05-0.95) as follows:

$$x_t = \frac{0.90(x_a - a)}{A - a} + 0.05 \quad \text{.....(5.1)}$$

The output results from the neural network are transformed back into their original value as follows.

$$x_a = \frac{(A-a)(x_t - 0.05)}{0.90} + a \quad \text{.....(5.2)}$$

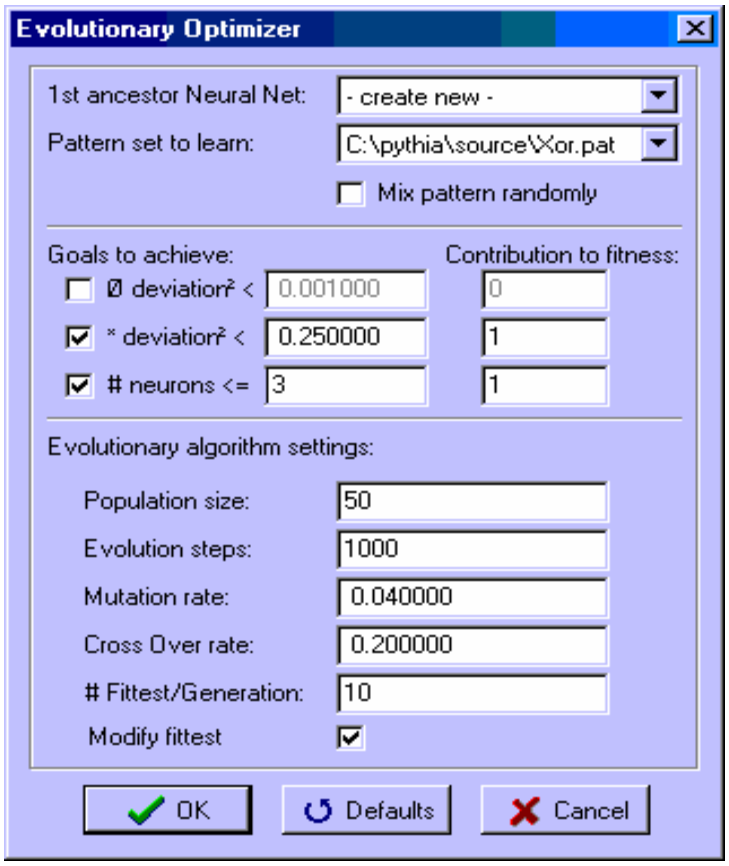


Figure (5.1): The Evolutionary Optimizer of the Pythia software, which is used as a tool for generating Neural Networks.

5.2.3 Software Used and Model Application:-

Many important issues arose while selecting software to be used to build and run the model. Those issues were; the simple user interaction with the model, managing the data files, executing the model, reviewing the results with simple output display features, the availability of the software, and the hardware that is compatible with the need to handle huge amount of data, run the model and the time required to perform these processes. Thus, Microsoft Excel Spreadsheet was selected to be utilized to build and run the model. The steps followed to construct the sheet are described bellow:

1- Insertion the input data sets with their order in columns.

2- Calculation the minimum and the maximum values for each column.

3- Normalization the input data.

4- Initialization of weights of the normalized input data.

5- Construction of the hidden layer No.1. In this step, summation of weighted input and the use of transfer function were carried for each neuron in this layer.

6- Repeat steps 3, 4 and 5 for all hidden layers and the output layer.

7- Insertion the actual output data set and calculation the corresponding minimum and maximum values.

8- Transformation back of the normalized output to their original values by using the maximum and minimum values of the actual output.

9- Calculation of the sum of squares and the efficiency.

10- Optimization of model weights using the solver algorithm. The optimization process was carried with different starting points.

5.3 Model Identification:-

From hydrological point of view the model was divided into two components each one describes a hydrological process namely: rainfall-runoff and channel routing processes. Figure (5.2) shows the system components. Each component consists of two blocks. Blocks (1) & (2) were responsible to simulate the rainfall-runoff process and provide forecasts of daily flows into Wed El-Heleiw and Kubur stations, respectively. Blocks (3) & (4) were responsible to simulate channel routing process and provide forecasts of daily inflows to Khashm El-Girba reservoir and Kilo_3 station, respectively.

Six years of lumped daily rainfall and streamflow data (1/1/1993 to 31/12/1998) were used in this study. Of these, 5 years (1/1/1993 to 31/12/1997) were selected for model calibration and the last year was used for model verification.

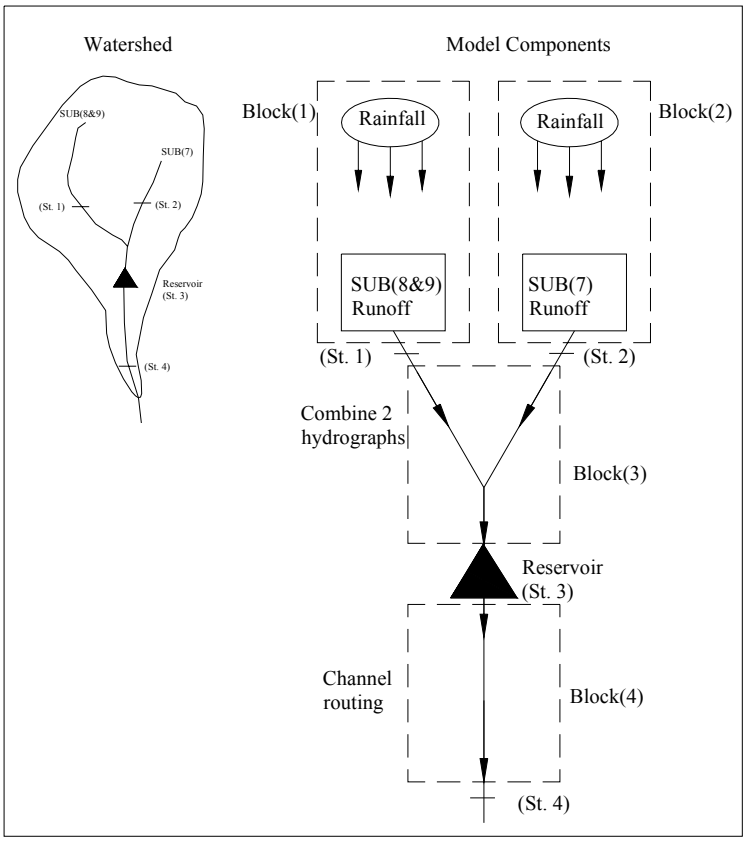


Figure (5.2): Schematic diagram shows the system modeled and its components.

5.3.1 Rainfall-Runoff Model:-

5.3.1.1 *Forecasts of Flows at Wed El-Heleiw Station (Block (1))*:-

5.3.1.1.1 *Model Identification*:-

Lumped daily rainfall time series was produced by averaging the lumped daily rainfall measurements of two rainfall stations namely: Gondar and Makale in the Ethiopian Highlands (see Figure (4.2)). The produced rainfall series accompanied with the streamflow data at Wed El-Heleiw station were used to build block (1).

First when using the TLM in the modeling process, a stepwise approach was followed in determining the optimum number of input lags (system memory length), in which the model was calibrated whenever an input lag is added. This process is stopped when the

addition of extra lags does not result in a significant improvement in the model performance with having stable parameters. The identified model with one-step lead streamflow forecast can be shown as follows:

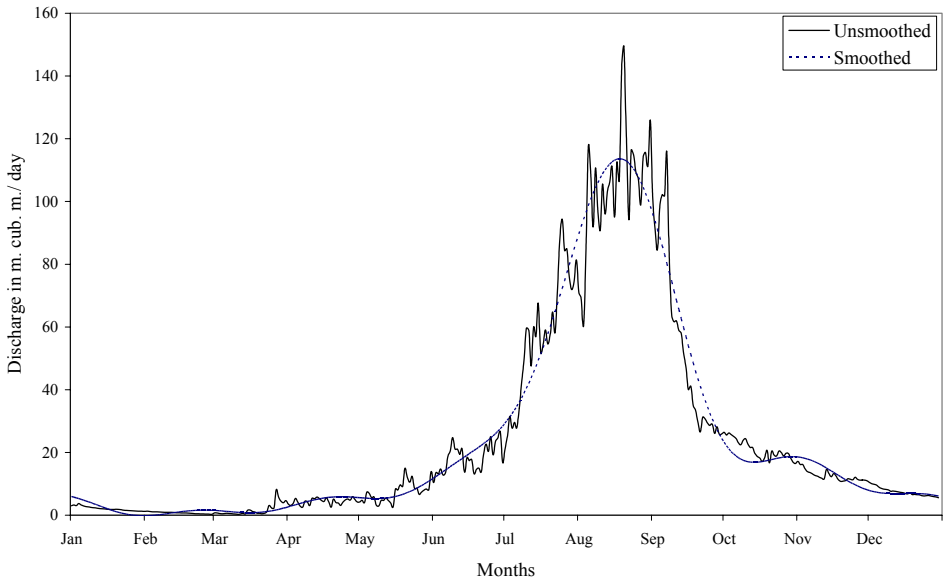
$$\hat{y}_t = \sum_{k=1}^{m} \hat{h}_k x_{t-k+1} \quad \text{.....(5.3)}$$

where $\hat{y}$ is the estimated discharge and $\hat{h}$ is the parameter estimates, and $m = 17$.

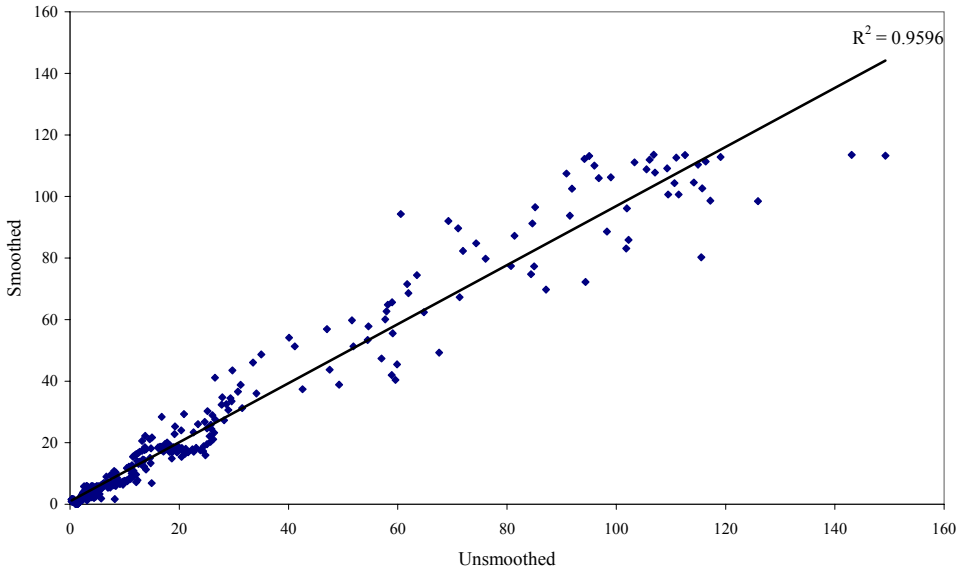
Second when using the LPM in the modeling process, the seasonal mean of flows for each date of the year (1 to 365) was obtained and the departures from the mean was calculated for both input and output variables over the calibration period. The mean values obtained for the output variable was then smoothed by the Fourier smoothing technique. The smoothing process was carried using different harmonics to find the most appropriate one. Six harmonics were found to be accepted since the smoothed curve fit both the low and high flows seasons satisfactory while accounting for more than 96% of the variance of the unsmoothed means see Figure(5.3). On the other hand, input output departure series were modeled using a combination of the TLM and the AR model, since the results of this process when using the TLM show a poor performance. The identified model of the input-output departure series can be shown by Equation (5.4), and the LPM can be viewed as a Modified LPM:

$$\hat{y}_t = \sum_{k=1}^{m} \hat{h}_k x_{t-k+1} + \sum_{k=1}^{n} \phi_k y_{t-k} \quad \text{.....(5.4)}$$

where $m = 4$, and $n = 1$.



(5.3a)



(5.3b)

Figure (5.3): Smoothed and unsmoothed seasonal means flow (a) hydrographs (b) scatter plot for the forecasting site no.1 (Wed El-Heleiw station).

Finally when using the ANNs model; based on the results of the traditional models (TLM, and MLPM) and on a systematic trail and error procedure the optimum number of

inputs and their lags were determined. A general representation of one-step lead streamflow forecast model can be shown as follows:

$$q_t = g_{non}(q_{t-1}, q_{t-2}, p_t, p_{t-1}, p_{t-2}, p_{t-3}) \quad \text{.....(5.5)}$$

where q_t and p_t are the streamflow and the daily rainfall, at time step t, respectively, and g_{non} represents a nonlinear transformation to be approximated by a neural network. The optimal network structure identified was (6, 10, 5, 1), with two hidden layers, the first hidden layer contains 10 neurons (60 weights) and the second layer contains 5 neurons (50 weights), bias weights were included in all hidden and output layers (the total number of parameters were 121) see Figure (5.4).

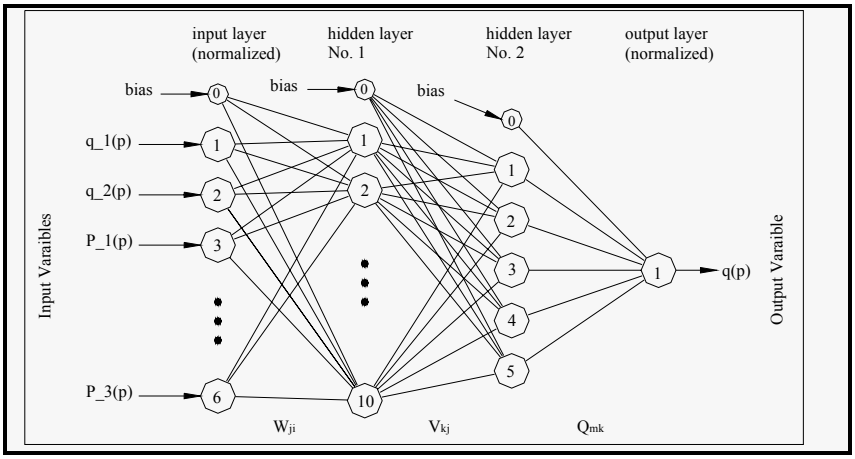


Figure (5.4): Schematic diagram shows the structure of the identified ANNs model for the forecasting site no.1 (Wed El-Heleiw station).

In the calibration process of the ANN model, initial values of weights, bias, and the factor of the Logistic function were set into random values. During calibration process and when the rate of increasing of efficiency slow down, the bias weights and the Logistic function factors were excluded from the optimization process. Two different activation functions were used; the Symmetric Logistic function (Equation (3.13)) used in both the first hidden layer and the output layer while the Logistic function (Equation (3.8)) used in the second hidden layer. The total time consumed to optimize the model parameters was about eight hours. Different starting points were tried at the calibration

process, and it all end with the same efficiency but with different values of optimum weights for each.

5.3.1.1.2 *Calibration and Verification Results*:-

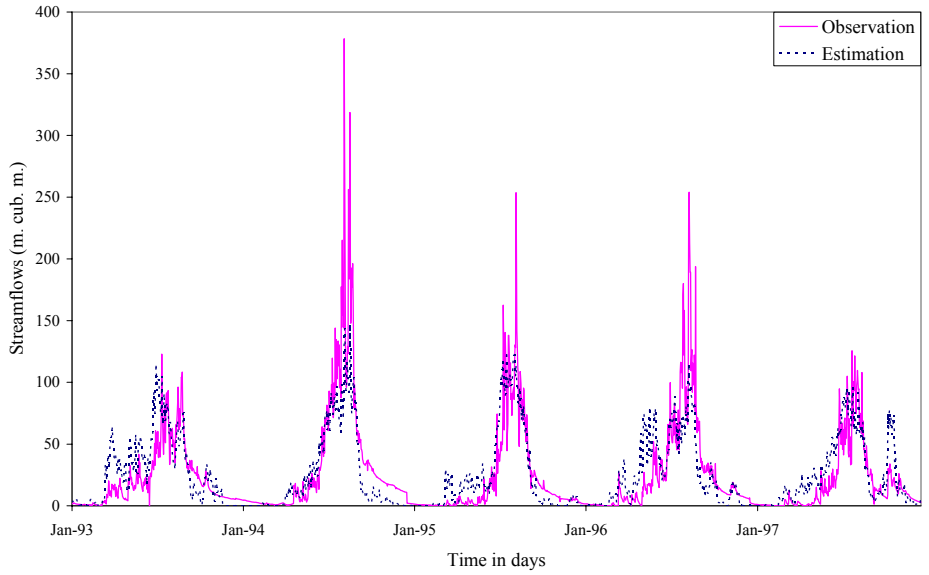
The calibration and validation period performances of all models are shown in Table (5.1), and details of their performance can be seen graphically in Figures (5.5) to (5.8).

Table (5.1a): Calibration and verification performance of the TLM, MLPM, and ANN at forecasting site no.1 (Wed El-Heleiw station) (Using 5 years (1993-1997) for calibration period and one year (1998) for verification period)

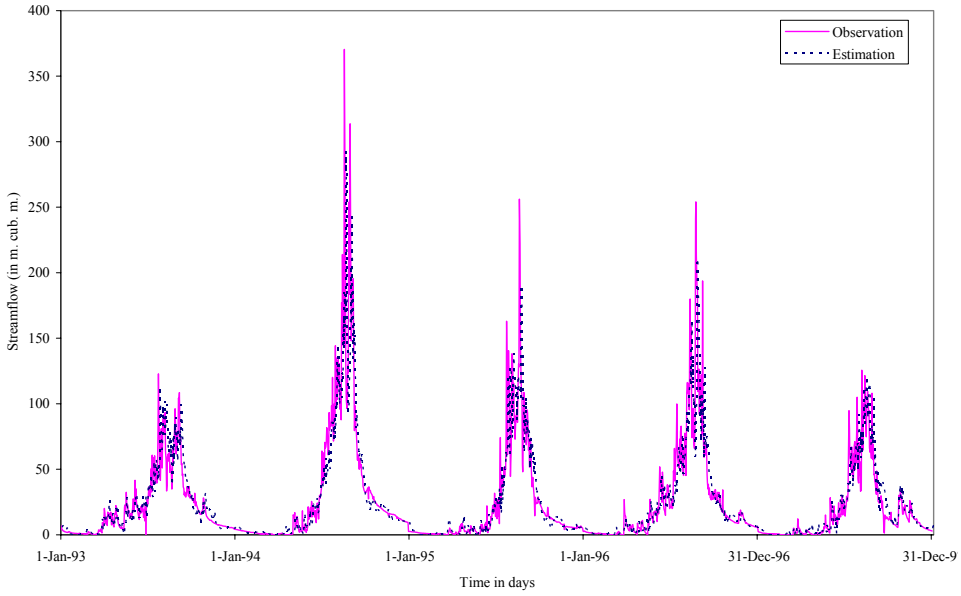
	Calibration			***Verification***		
	TLM	**MLPM**	**ANN**	**TLM**	**MLPM**	**ANN**
MIN. E	-78.17	-62.08	-131.70	-96.29	-150.65	-505.02
MAX.E	233.32	200.28	71.20	793.15	579.64	313.86
SSE(*10^3)	970.789	345.118	234.397	265.7680	101.175	892.814
RMSE	23.16	13.76	11.34	88.67	53.69	48.73
R^2	0.63	0.87	0.91	0.54	0.79	0.81
Eff (%)	63.12	86.76	91.03	40.80	77.64	80.40
Bias (%)	4.26	0.45	5.17	-38.31	-14.35	9.01
%MF	-62.99	-54.07	-11.91	-87.20	-63.73	-55.52

Table (5.1b): Calibration and verification performance of the TLM, MLPM, and ANN at forecasting site no.1 (Wed El-Heleiw station) (Using 4 years (1993-1996) for calibration and one year (1997) for verification).

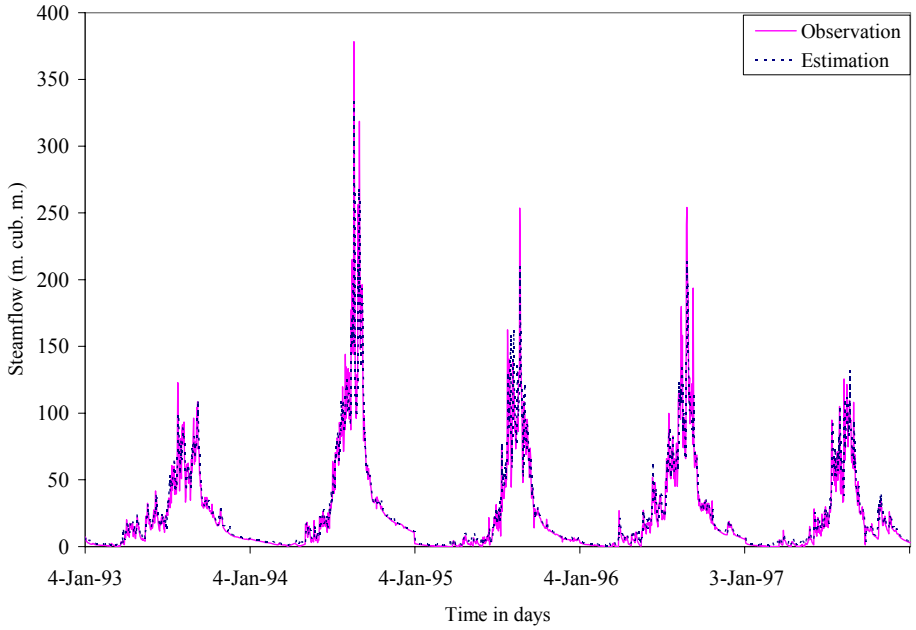
	Calibration			***Verification***		
	TLM	**MLPM**	**ANN**	**TLM**	**MLPM**	**ANN**
MIN. E	-81.53	-62.08	-143.81	-52.23	-48.39	-88.27
MAX.E	234.38	200.28	73.73	51.35	59.55	26.59
SSE(*10^3)	804.555	306.462	253.152	124.021	49.268	36.609
RMSE	23.60	14.498	11.78	18.88	11.85	9.87
R^2	0.66	0.87	0.89	0.65	0.83	0.8729
Eff (%)	66.19%	86.93	89.22	48.00	79.37	84.88
Bias (%)	1.17%	-4.38	6.98	43.87	15.20	-8.55
%MF	-61.96	-54.07	-11.908	-7.26	1.62	-24.03


Observation
Estimation
Streamflows (m. cub. m.)
400
350
300
250
200
150
100
50
0
Jan-93
Jan-94
Jan-95
Jan-96
Jan-97
Time in days

(5.5a)

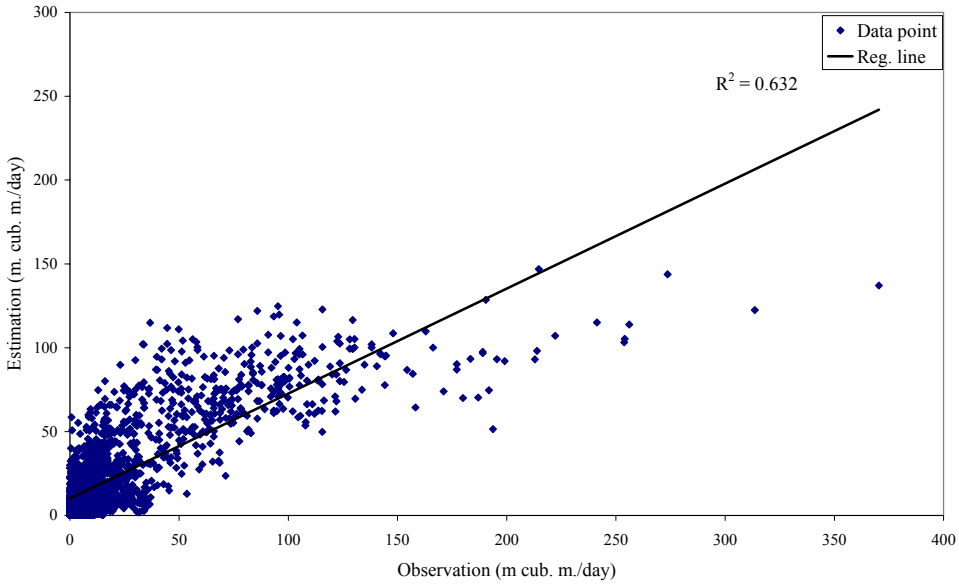

Observation
Estimation
Streamflow (in m. cub. m.)
400
350
300
250
200
150
100
50
0
1-Jan-93
1-Jan-94
1-Jan-95
1-Jan-96
31-Dec-96
31-Dec-9
Time in days

(5.5b)

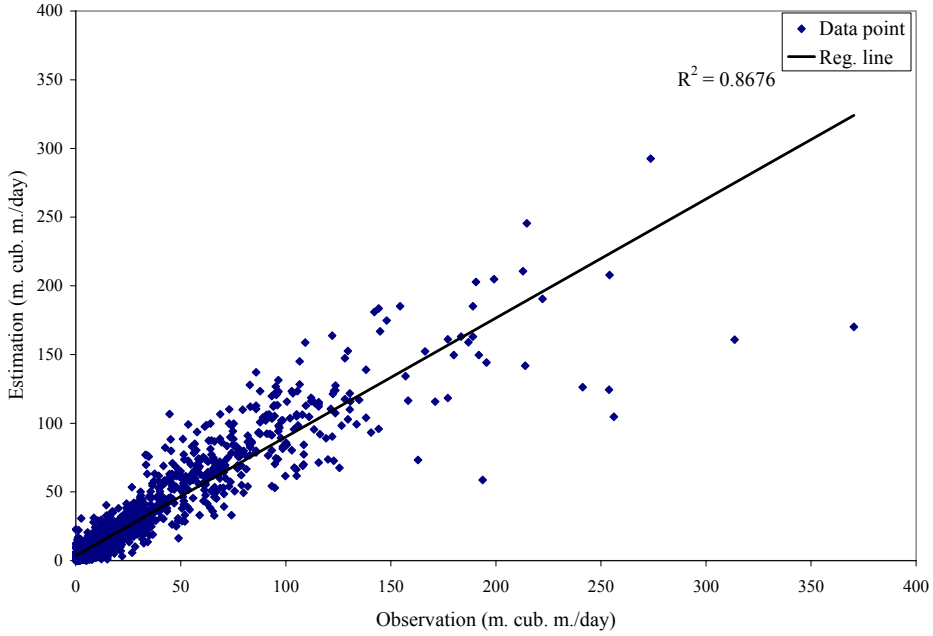


(5.5c)

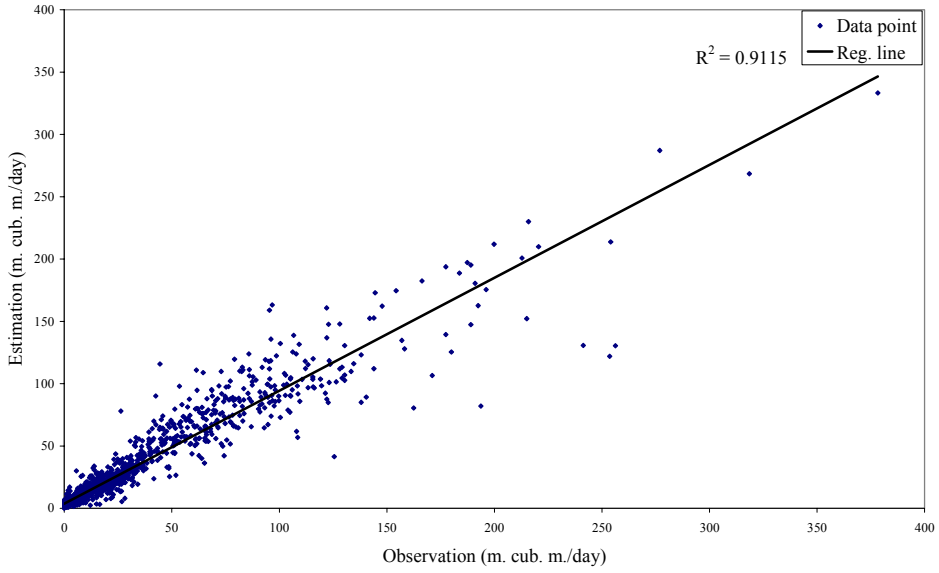
Figure (5.5): Observed and estimated daily streamflow hydrographs during calibration period at forecasting site no.1 (Wed El-Heleiw station) using: (a) TLM, (b) MLPM, and (c) ANNs.



(5.6a)

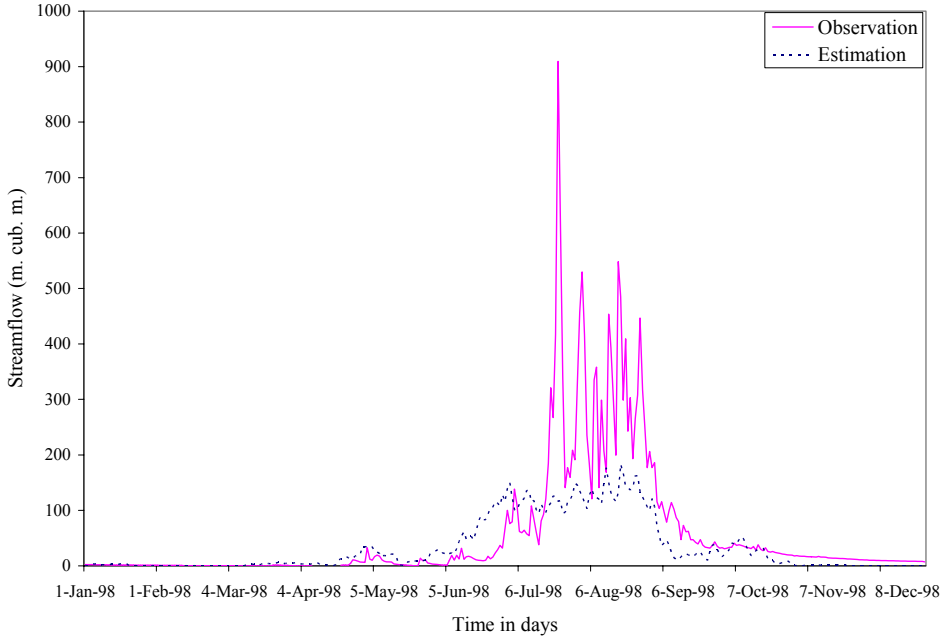


(5.6b)

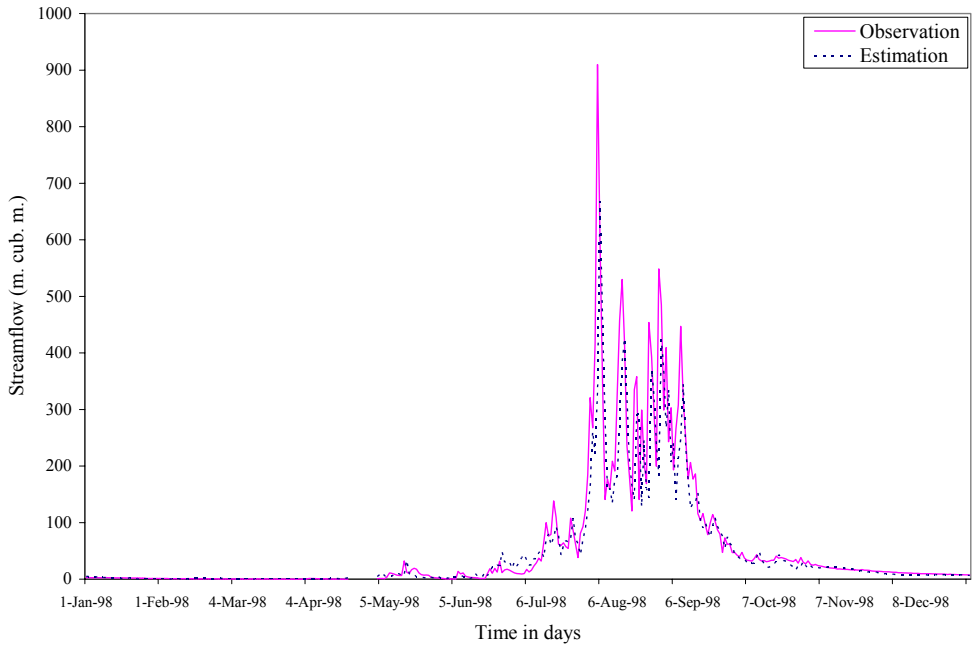


(5.6c)

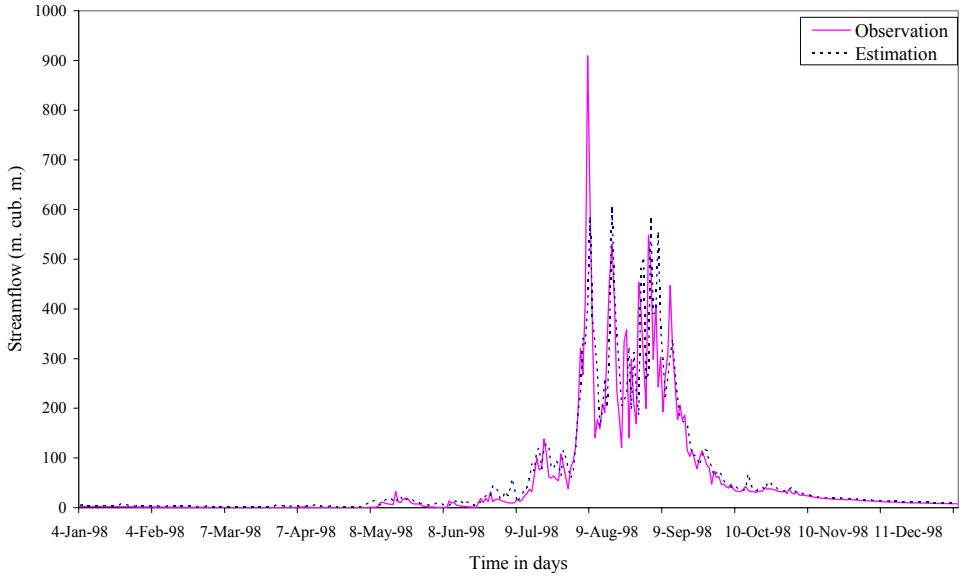
Figure (5.6): Observed and estimated daily streamflow scatterplots during calibration period at forecasting site no.1 (Wed El-Heleiw station) using: (a) TLM, (b) MLPM, and (c) ANNs.


Observation
Estimation
Streamflow (m. cub. m.)
1000
900
800
700
600
500
400
300
200
100
0
1-Jan-98
1-Feb-98
4-Mar-98
4-Apr-98
5-May-98
5-Jun-98
6-Jul-98
6-Aug-98
6-Sep-98
7-Oct-98
7-Nov-98
8-Dec-98
Time in days

(5.7a)

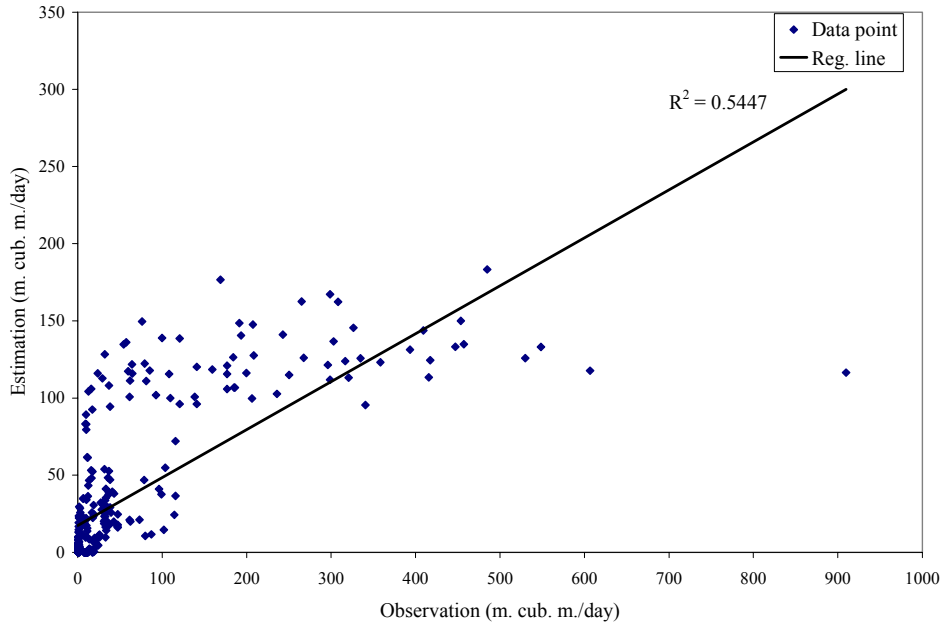

Observation
Estimation
Streamflow (m. cub. m.)
1000
900
800
700
600
500
400
300
200
100
0
1-Jan-98
1-Feb-98
4-Mar-98
4-Apr-98
5-May-98
5-Jun-98
6-Jul-98
6-Aug-98
6-Sep-98
7-Oct-98
7-Nov-98
8-Dec-98
Time in days

(5.7b)

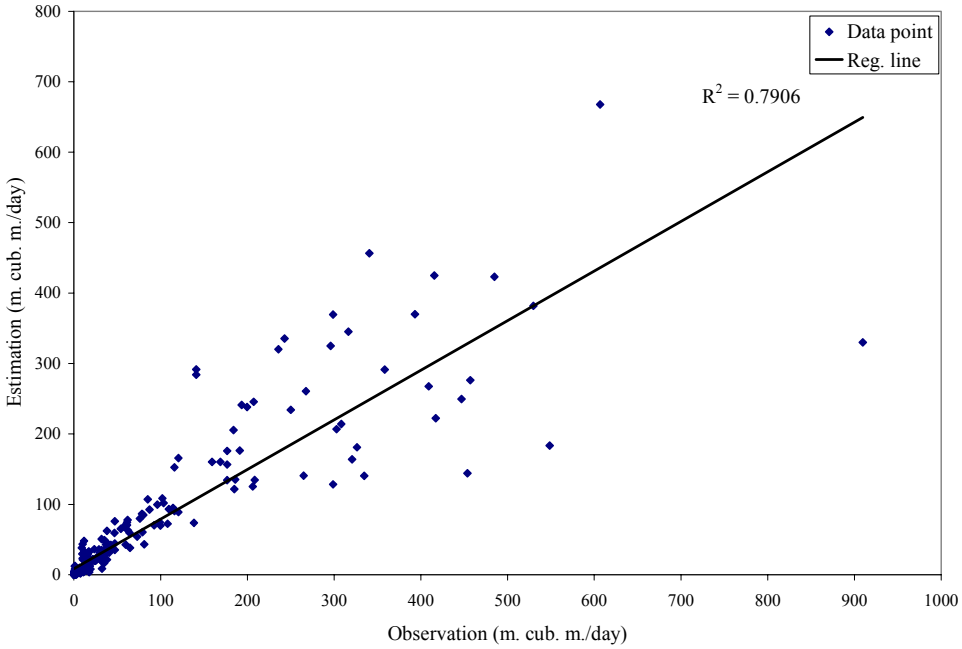


(5.7c)

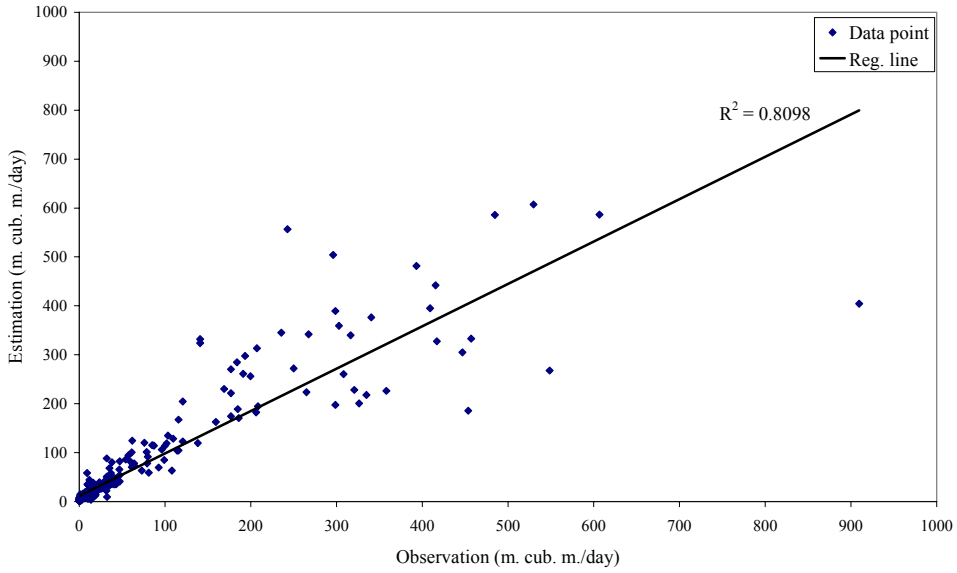
Figure (5.7): Observed and estimated daily streamflow hydrographs during verification period at forecasting site no.1 (Wed El-Heleiw station) using: (a) TLM, (b) MLPM, and (c) ANNs.



(5.8a)



(5.8b)



(5.8c)

Figure (5.8): Observed and estimated daily streamflow scatterplots during verification period at forecasting site no.1 (Wed El-Heleiw station) using: (a) TLM, (b) MLPM, and (c) ANNs.

5.3.1.2 *Forecast of Flows at Kubur Station Model*:-

5.3.1.2.1 *Model Identification*:-

Lumped daily rainfall time series was produced by averaging lumped daily rainfall measurements of two rainfall station in the Ethiopian Highlands namely: Godar and Bahar_Dar (Figure (4.2)). The produced series accompanied with the streamflow data at Kubur station were used to build block (2).

First, the modeling process was carried using the TLM, the optimum number of input lags was determined using the same procedure followed in block (1). The identified model with one-step lead streamflow forecast can be shown as it was by Equation (5.3.), where m was found to be 34.

Second, the AR (p) model was used in the modeling process, the order p of the autoregressive process was determined by calculating and plotting the partial autocorrelations as shown in Figure (5.9).The identified model with one-step ahead streamflow forecast can be shown as it was by Equation (3.41), where p was found to be 3.

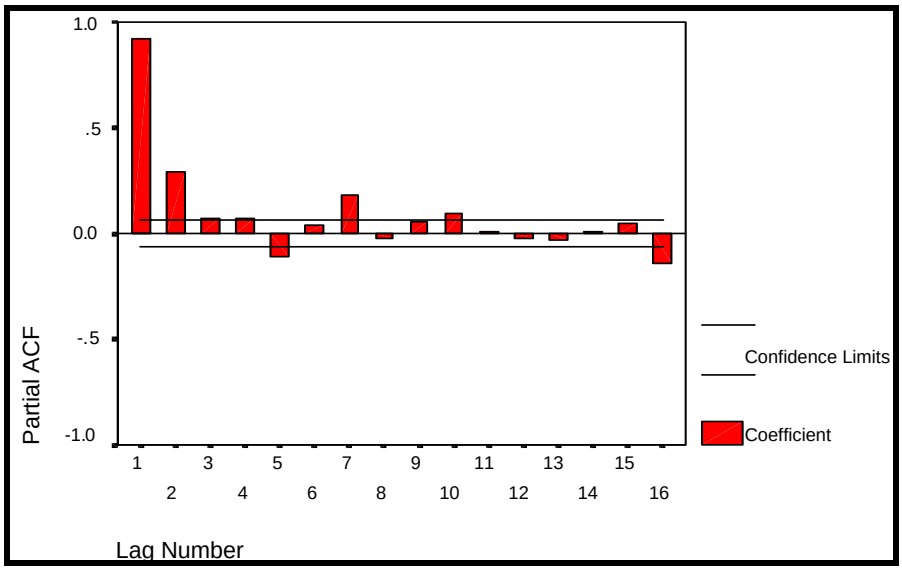


Figure (5.9): Plot of the Partial Autocorrelations vs. the number of lags for Kubur Station.

Third, a combination of the AR(p) model and the TLM (AR-TLM) was used in the model building process. The model parameters were optimized using a stepwise approach, tacking into account the stability of the model parameters. The identified model with one-step lead streamflow forecast can be shown as it was by Equation (5.4), where k (corresponding to m) = 2, and $n = 1$.

Four, when using the LPM in the modeling process, the seasonal mean of flows for each date of the year (1 to 365) was obtained and the departures from the mean was calculated for both input and output variables over the calibration period. The mean values obtained for the output variable was then smoothed by the Fourier smoothing technique. It is worth mentioning that the streamflow covers only about five months. The smoothing of the mean values was carried using different harmonics and nine harmonics were found to be accepted (Figure (5.10)). On the other hand, input output departure series were modeled using a combination of the TLM and the AR model, since the results of this process when using the TLM shows poor performance. The identified model of the input-output departure series can be shown Equation (5.4), where k (corresponding to m) = 2, and $n = 1$, and the LPM can be viewed as a Modified LPM.

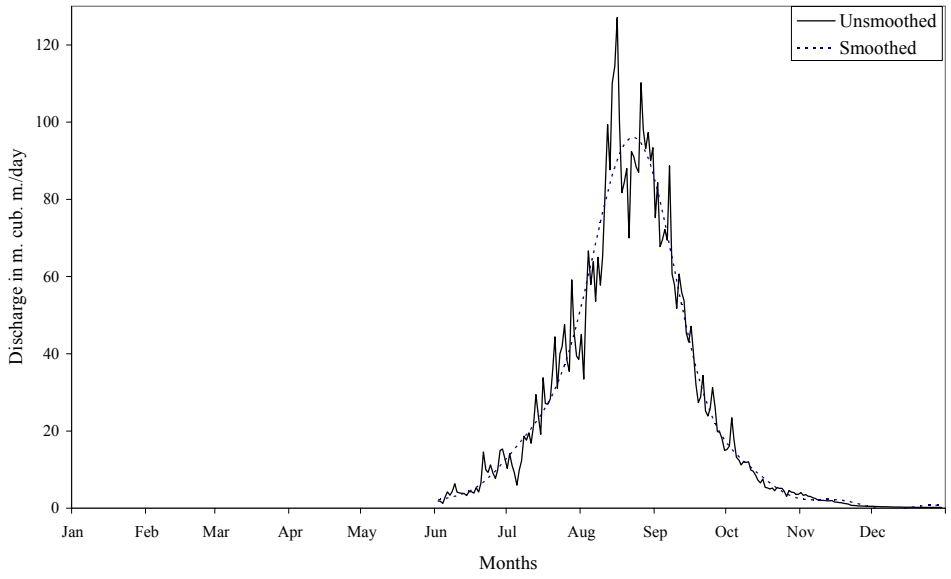


Figure (5.10): Smoothed and unsmoothed seasonal means flow hydrographs at forecasting site no.2 (Kubur station).

Finally when using ANNs, the optimum number of model inputs in this block was determined using the same procedure that followed in block (1). A general representation of one-step-ahead streamflow forecast model can be shown as follows:

$$q_t = g_{non}(q_{t-1}, q_{t-2}, q_{t-3}, p_t, p_{t-1}, p_{t-2}, p_{t-3}) \quad \text{.....(5.6)}$$

where q_t and p_t were the streamflow and the daily rainfall, at time step t, respectively, and g_{non} represents a nonlinear transformation to be approximated by a neural network.

The optimum network structure identified was (7,9,7,1), with two hidden layers, the first layer contains 9 neurons (63 weights) and the second layer with 7 neurons(63 weights), bias weights were included in all hidden and output layers(the total number of model parameters were 139) (Figure (5.11)).

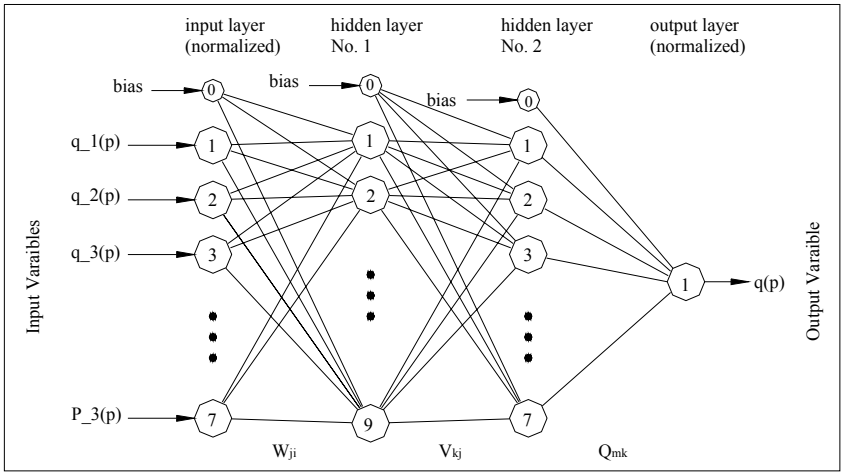


Figure (5.11): Schematic diagram shows the structure of the identified ANNs model at forecasting site no.2 (Kubur station).

In the calibration process of the ANN model, initial values of weights, bias, and the factor of the Logistic function were set into random values. During calibration process and when the rate of increasing of efficiency slow down, the bias weights and the Logistic function factors were excluded from the optimization process. Two different activation functions were used, the Symmetric Logistic function (Equation (3.13)) was used in the first hidden layer while in the second hidden layer the Logistic function (Equation (3.8)) was used and in the output layer the Symmetric Logistic function was used. The total time consumed to optimize the model parameters was about nine hours. It is worth mentioning that different starting points were tried during the calibration process. Although, different values were found in all trials for each weight, all trials end up with the same efficiency.

5.3.1.2.2 *Calibration and Verification Results*:-

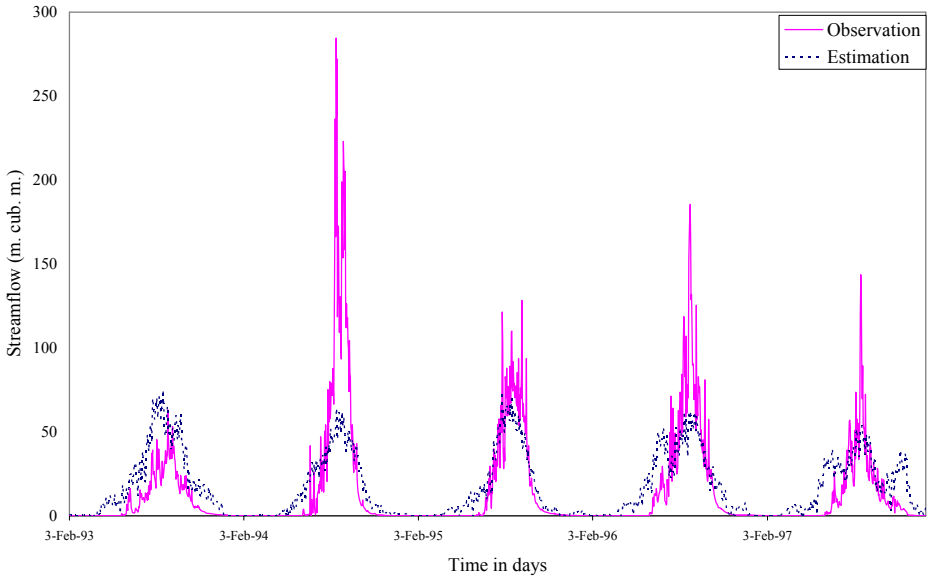
The calibration and validation period performances of all models are shown in Table (5.2), and details of their performance can be seen graphically in Figures (5.12) to (5.15).

Table (5.2a): Calibration and verification performance of the TLM, AR, AR-TLM, MLPM and ANN at the forecasting site no.2 (Kubur station) (Using 5 years (1993-1997) for calibration period and one year (1998) for verification period).

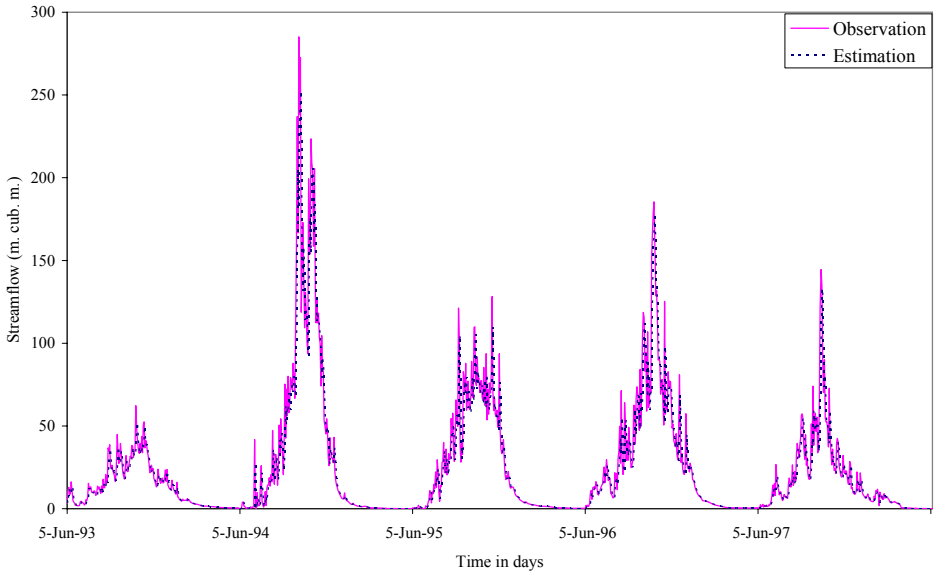
	Calibration					*Verification*				
	TLM	**AR**	**AR-TLM**	**MLPM**	**ANN**	**TLM**	**AR**	**AR-TLM**	**MLPM**	**ANN**
MIN. E	-224.18	-133.44	-127.98	-114.26	-88.71	-164.33	-74.90	-70.80	-67.11	-77.71
MAX.E	52.42	107.00	123.68	119.26	41.26	191.57	108.36	116.73	108.74	90.16
SSE(*10^3)	874.192	196.718	205.210	197.948	128.843	401.247	89.907	106.838	119.804	86.548
RMSE	22.08	13.64	13.93	13.70	10.97	44.57	20.89	22.56	22.82	19.93
R^2	0.51	0.87	0.86	0.87	0.91	0.37	0.85	0.83	0.82	0.86
Eff (%)	62.69	86.57	85.99	86.45	91.18	32.98	85.31	82.76	81.46	85.54
Bias (%)	18.13	-3.67	-0.79	2.52	-1.17	5.37	-3.72	-2.79	8.06	-5.27
%MF	-78.77	-36.27	-43.37	-41.90	-17.87	-70.84	-33.66	-45.58	-36.81	-17.61

Table (5.2b): Calibration and verification performance of the TLM, AR, AR-TLM, MLPM and ANN at the forecasting site no.2 (Kubur station), (Using 4 years (1993-1996) for calibration period and one year (1997) for verification period).

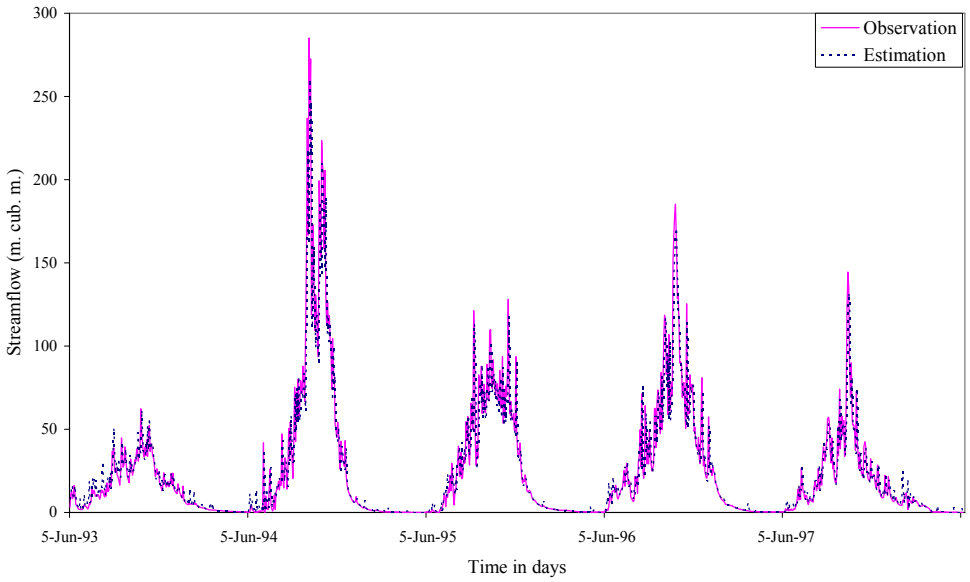
	Calibration					*Verification*				
	TLM	**AR**	**AR-TLM**	**MLPM**	**ANN**	**TLM**	**AR**	**AR-TLM**	**MLPM**	**ANN**
MIN. E	-224.18	-133.10	-128.19	-101.33	-83.78	-164.33	-43.11	-35.89	-48.47	-38.68
MAX.E	52.42	106.69	123.30	139.40	45.11	191.57	67.41	53.50	68.52	33.25
SSE(*10^3)	786.05	175.57	140.43	174.82	113.45	401.25	212.45	188.32	29853.30	136.24
RMSE	23.46	14.41	15.26	14.40	10.30	44.57	10.16	9.47	12.16	7.76
R^2	0.52	0.87	0.86	0.87	0.9152	0.37	0.81	0.83	0.92	0.8742
Eff (%)	59.23	86.89	86.12	86.93	91.51	32.98	80.76	83.10	72.31	87.66
Bias (%)	10.75	-3.52	-0.49	0.43	0.70	5.37	-3.50	2.18	8.34	-2.39
%MF	-78.77	-36.27	-43.37	-41.90	-17.87	-70.84	7.15	-12.11	-31.61	33.01


Observation
Estimation
Streamflow (m. cub. m.)
300
250
200
150
100
50
0
3-Feb-93
3-Feb-94
3-Feb-95
3-Feb-96
3-Feb-97
Time in days

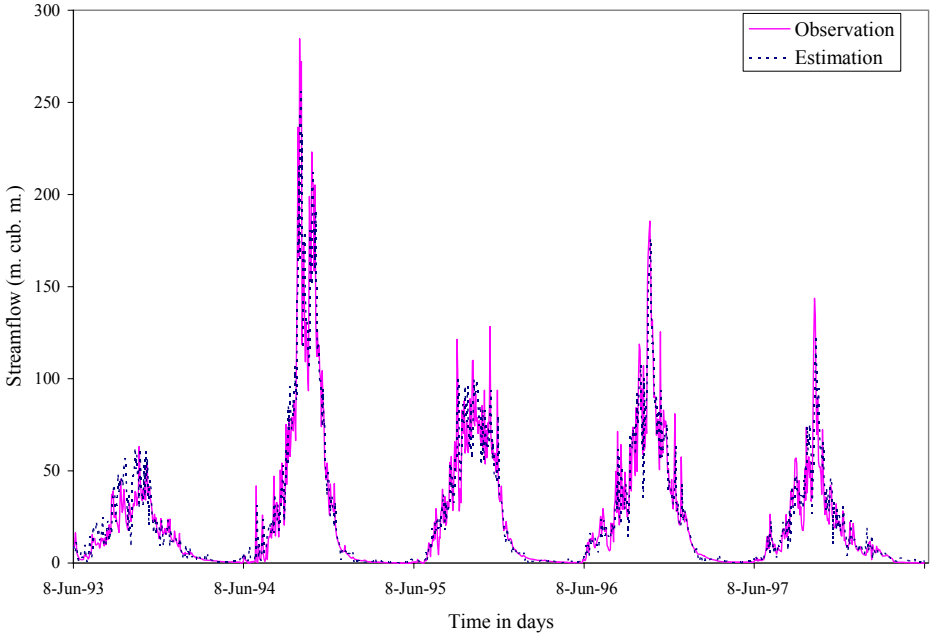
(5.12a)


Observation
Estimation
Streamflow (m. cub. m.)
300
250
200
150
100
50
0
5-Jun-93
5-Jun-94
5-Jun-95
5-Jun-96
5-Jun-97
Time in days

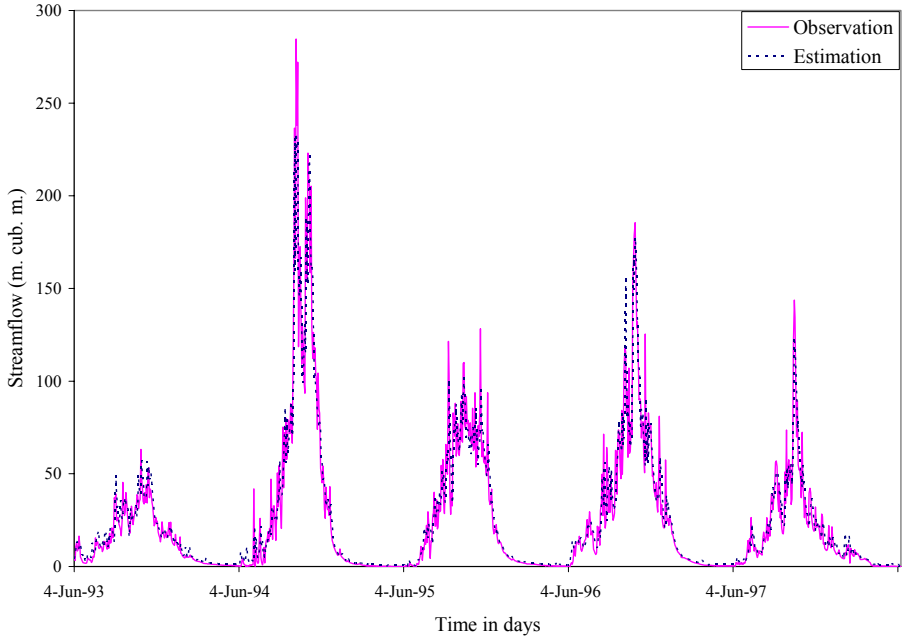
(5.12b)


Observation
Estimation
Streamflow (m. cub. m.)
300
250
200
150
100
50
0
5-Jun-93
5-Jun-94
5-Jun-95
5-Jun-96
5-Jun-97
Time in days

(5.12c)

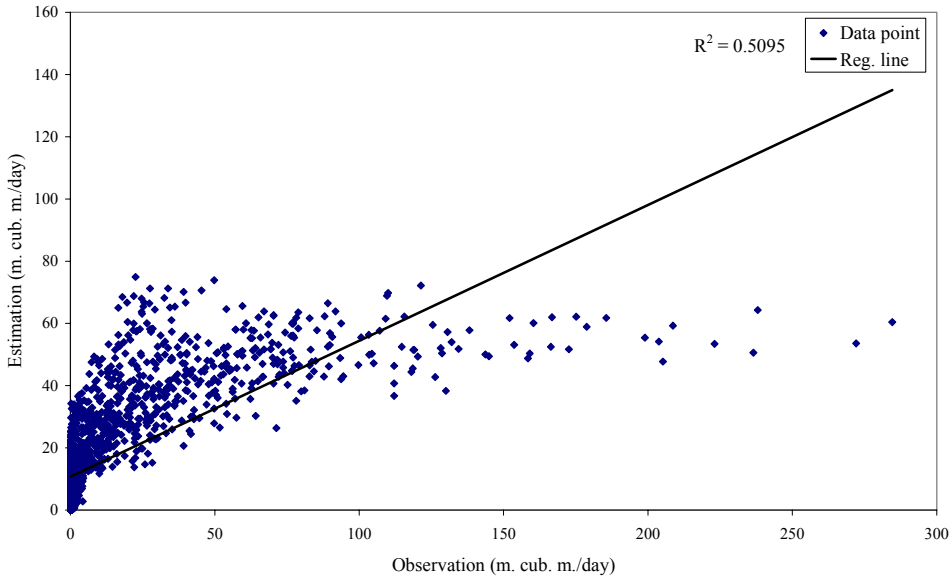

Observation
Estimation
Streamflow (m. cub. m.)
300
250
200
150
100
50
0
8-Jun-93
8-Jun-94
8-Jun-95
8-Jun-96
8-Jun-97
Time in days

(5.12d)

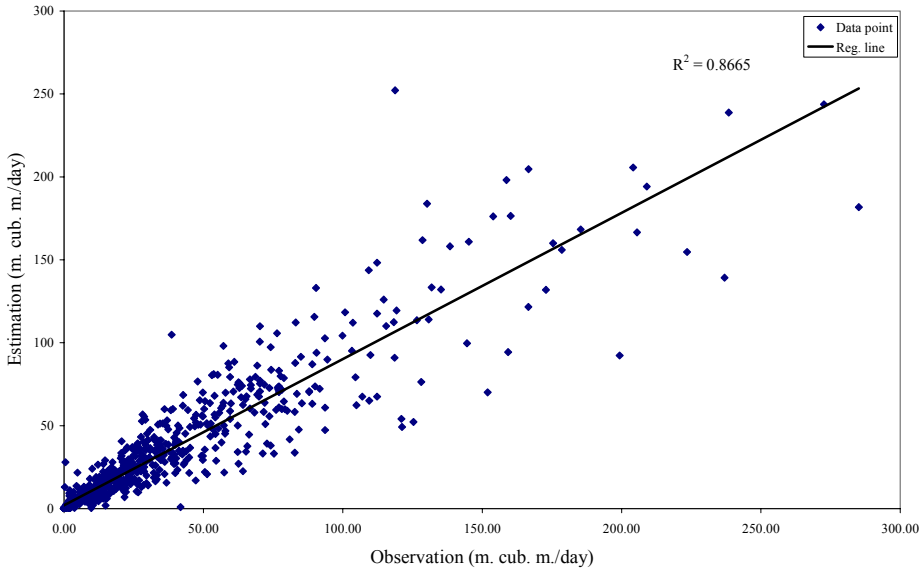


(5.12e)

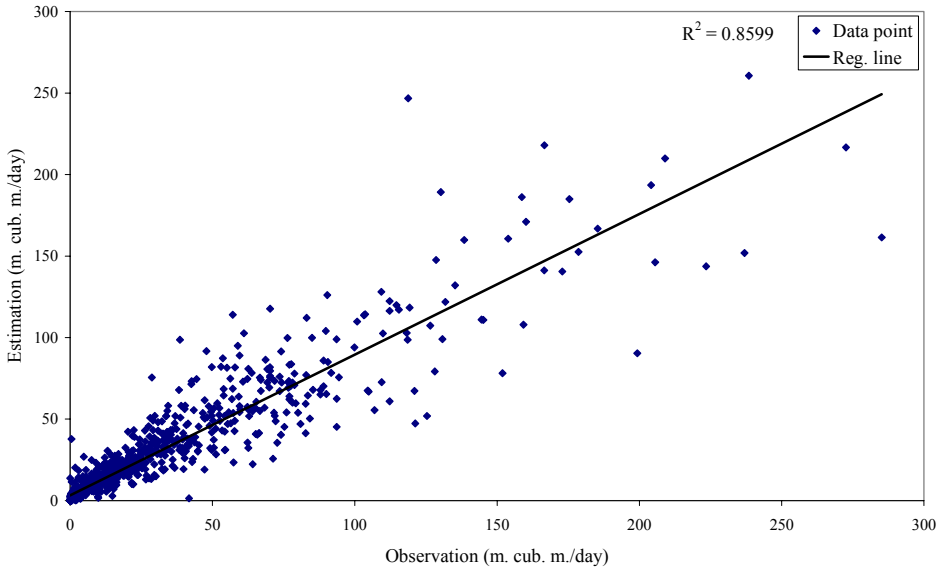
Figure (5.12): Observed and estimated daily streamflow hydrographs during calibration period at forecasting site no.2 (Kubur station) using: (a) TLM, (b) AR, (c) AR-TLM, (d) MLPM, and (e) ANNs.



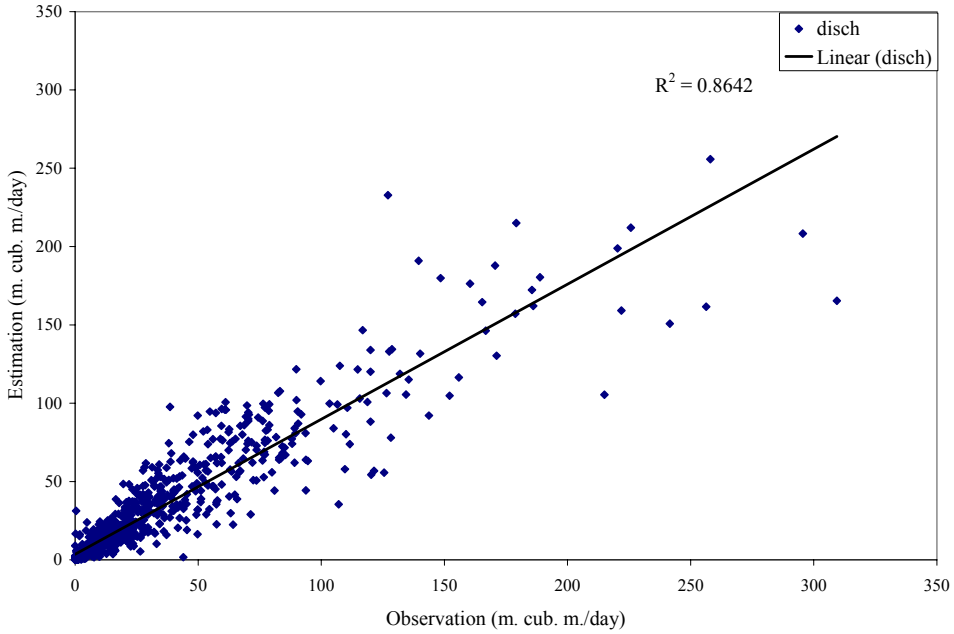
(5.13a)


Data point
Reg. line
$R^2 = 0.8665$
Estimation (m. cub. m./day)
Observation (m. cub. m./day)
0
50
100
150
200
250
300
0.00
50.00
100.00
150.00
200.00
250.00
300.00

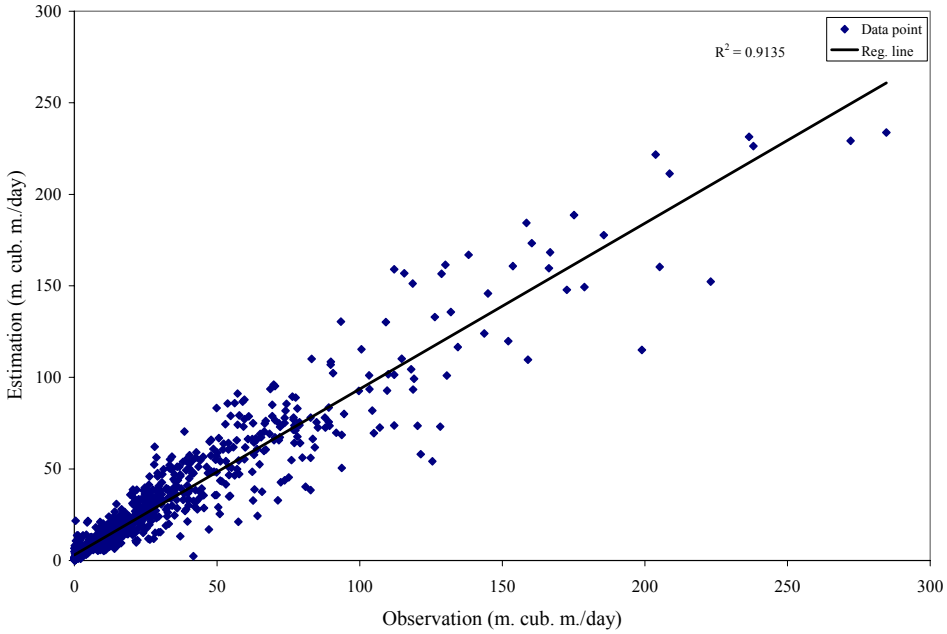
(5.13b)


Data point
Reg. line
$R^2 = 0.8599$
Estimation (m. cub. m./day)
Observation (m. cub. m./day)
0
50
100
150
200
250
300

(5.13c)

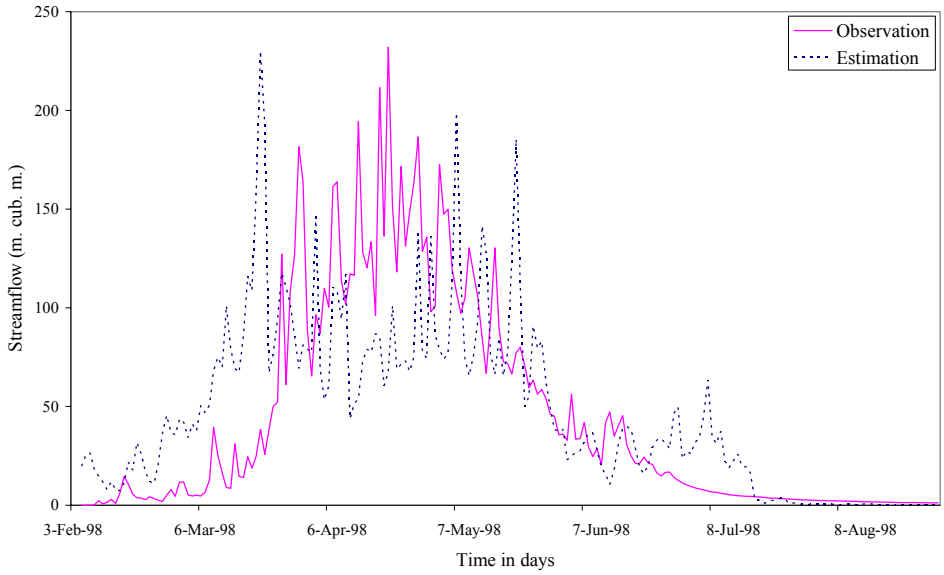


(5.13d)

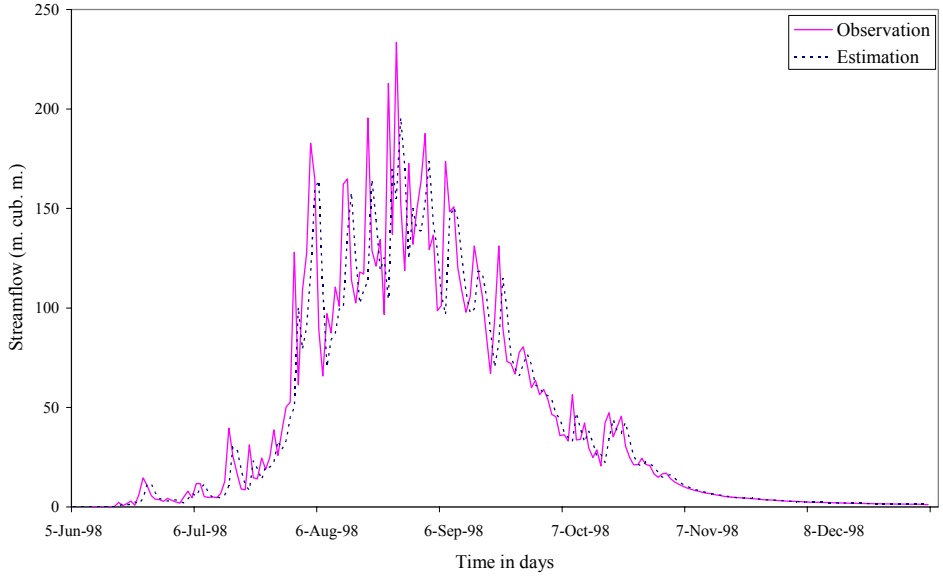


(5.13e)

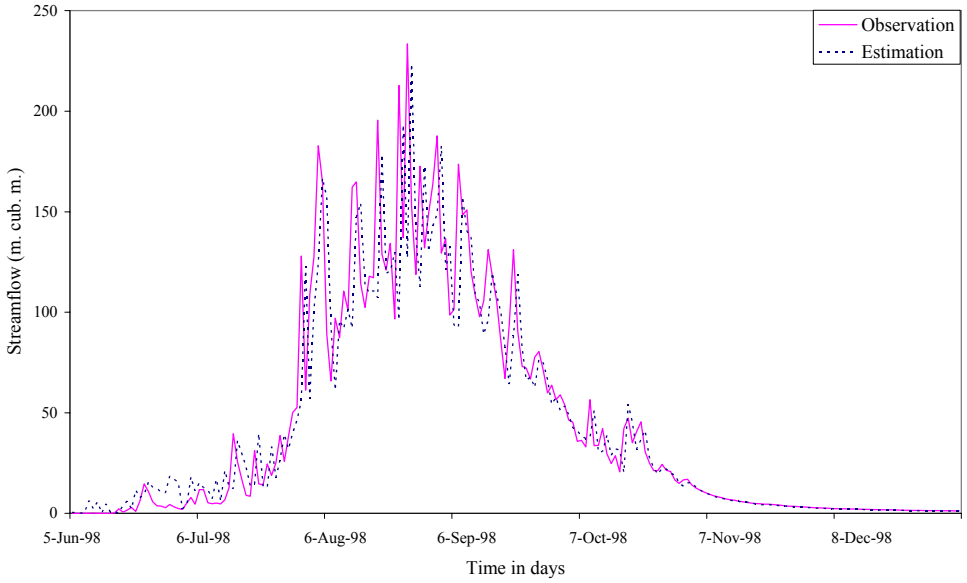
Figure (5.13): Observed and estimated daily streamflow scatterplots during calibration period at forecasting site no.2 (Kubur station) using: (a) TLM, (b) AR, (c) AR-TLM, (d) MLPM, and (e) ANNs.


Observation
Estimation
Streamflow (m. cub. m.)
250
200
150
100
50
0
3-Feb-98
6-Mar-98
6-Apr-98
7-May-98
7-Jun-98
8-Jul-98
8-Aug-98
Time in days

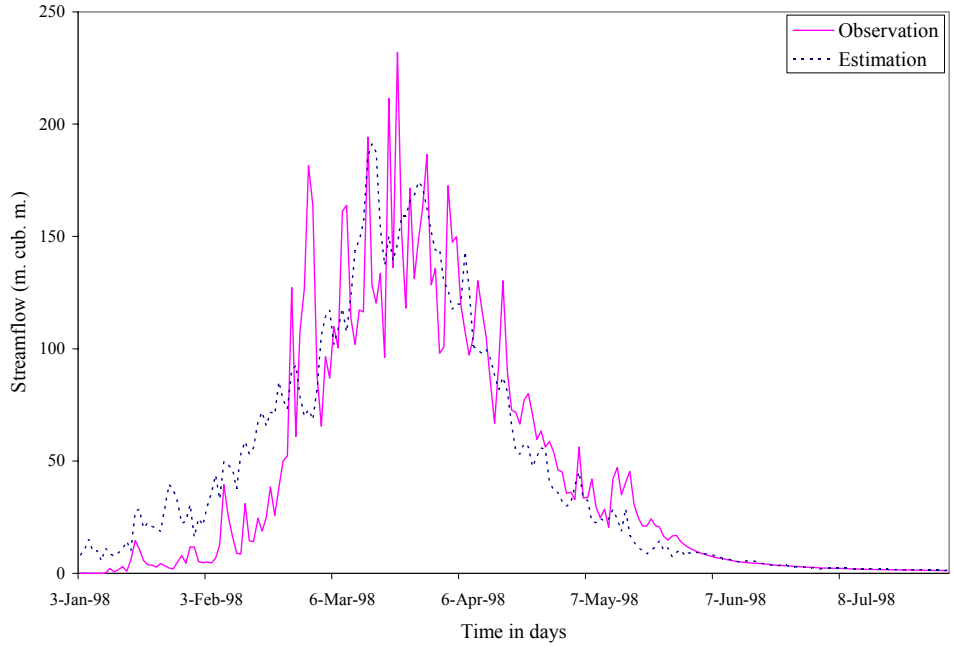
(5.14a)


Observation
Estimation
Streamflow (m. cub. m.)
250
200
150
100
50
0
5-Jun-98
6-Jul-98
6-Aug-98
6-Sep-98
7-Oct-98
7-Nov-98
8-Dec-98
Time in days

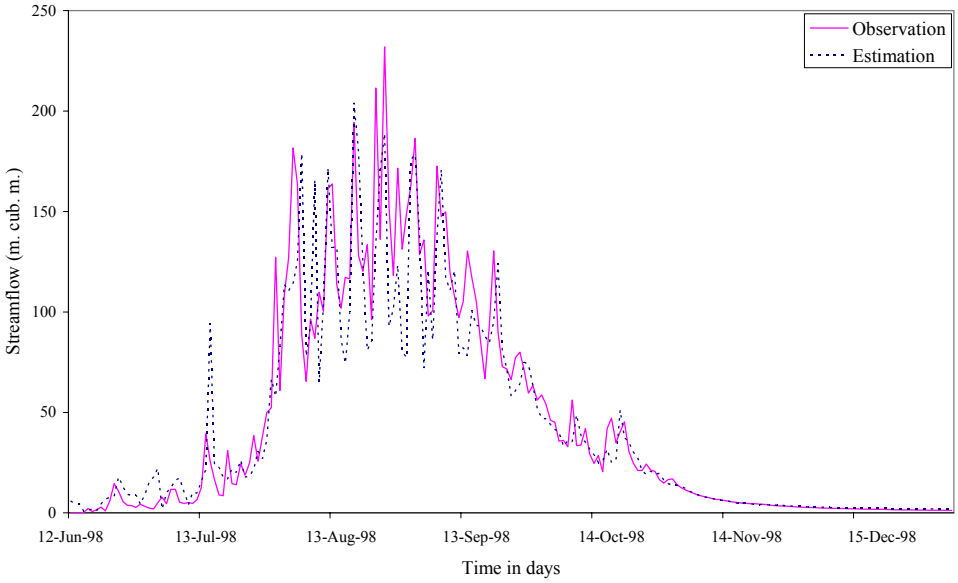
(5.14b)


Observation
Estimation
Streamflow (m. cub. m.)
250
200
150
100
50
0
5-Jun-98
6-Jul-98
6-Aug-98
6-Sep-98
7-Oct-98
7-Nov-98
8-Dec-98
Time in days

(5.14c)

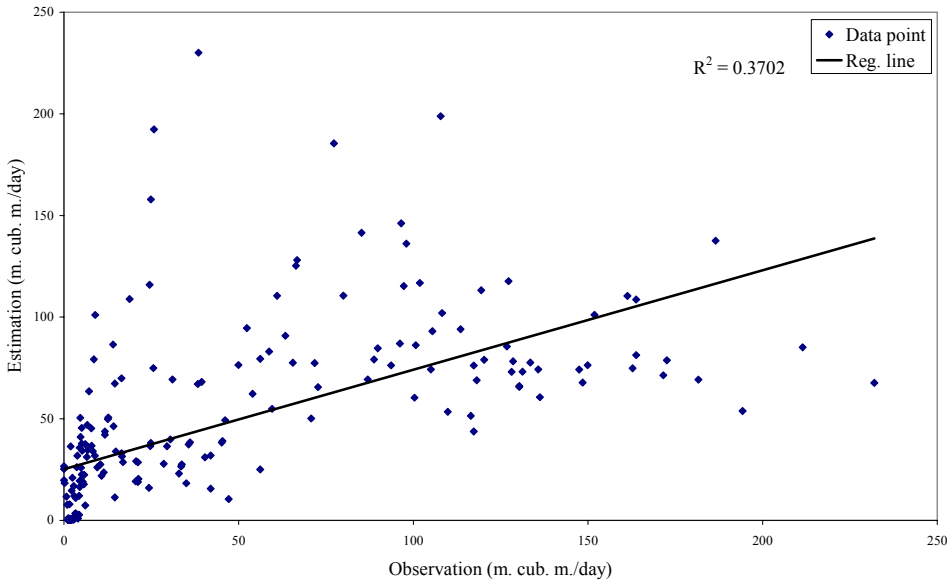

Observation
Estimation
Streamflow (m. cub. m.)
250
200
150
100
50
0
3-Jan-98
3-Feb-98
6-Mar-98
6-Apr-98
7-May-98
7-Jun-98
8-Jul-98
Time in days

(5.14d)

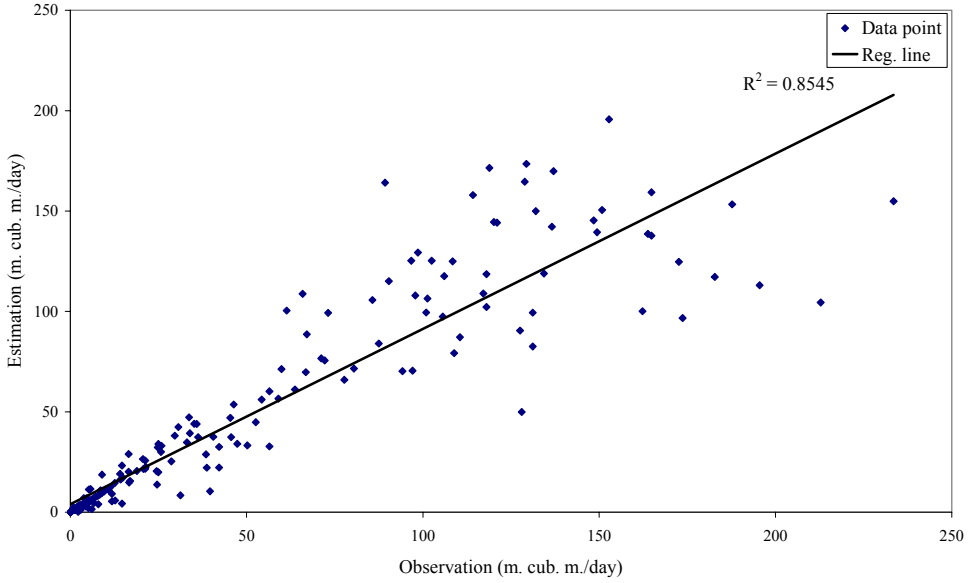


(5.14e)

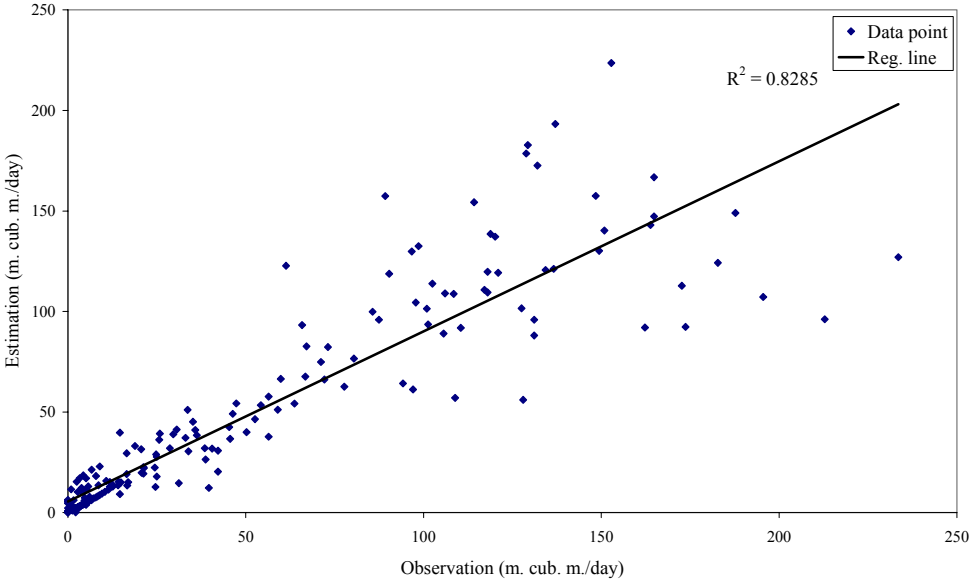
Figure (5.14): Observed and estimated daily streamflow hydrographs during verification period at forecasting site no.2 (Kubur station) using: (a) TLM, (b) AR, (c) AR-TLM, (d) MLPM, and (e) ANNs.



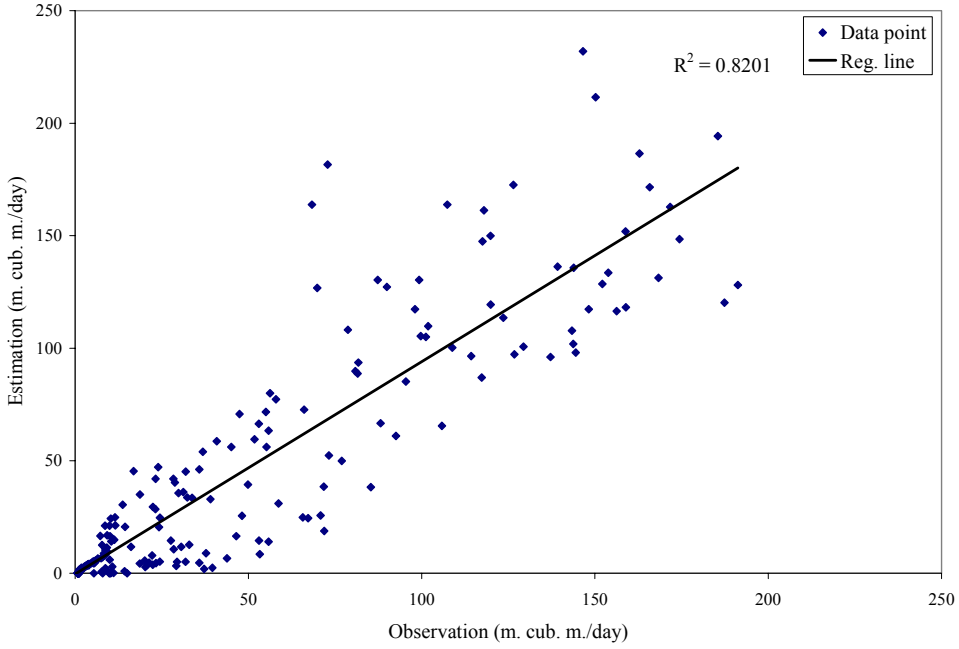
(5.15a)


Data point
Reg. line
R^2 = 0.8545
Estimation (m. cub. m./day)
Observation (m. cub. m./day)
0
50
100
150
200
250

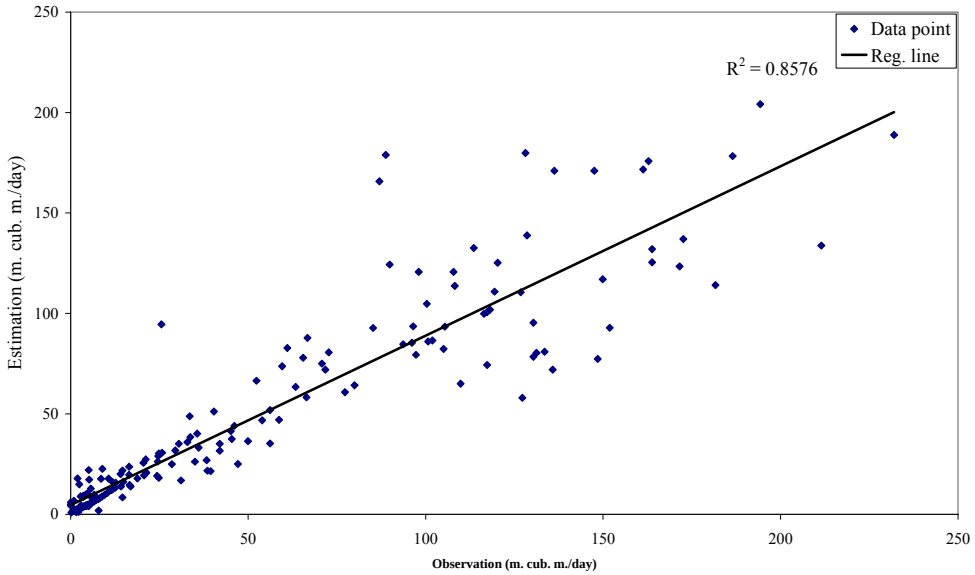
(5.15b)


Data point
Reg. line
R^2 = 0.8285
Estimation (m. cub. m./day)
Observation (m. cub. m./day)
0
50
100
150
200
250

(5.15c)



(5.15d)



(5.15e)

Figure (5.15): Observed and estimated daily streamflow scatterplots during verification period at forecasting site no.2 (Kubur station) using: (a) TLM, (b) AR, (c) AR-TLM, (d) MLPM, and (e) ANNs.

5.3.2 Channel Routing Model:-

5.3.2.1 *Forecast of Inflows into Khashm El-Girba Reservoir*:-

5.3.2.1.1 *Model Identification*:-

Streamflow data of Wed El-Heleiw and Kubur Stations (about 500 m and 443 m upstream of Khashm El-Girba reservoir respectively) (as an input variables) accompanied with the daily inflows into Khashm El-Girba reservoir (as an output variable) were used to build block (3).

First when using the TLM, the optimum number of input lags was determined using stepwise approach, by calibrating the model each time an input lag interred. The resulting models were assessed considering their performance and the parameters stability. The identified model can be shown by Equation (5.7)

$$\hat{y}_t = \sum_{k=1}^{m} \hat{h}_k x_{t-k+1} + \sum_{k=1}^{n} \hat{a}_k z_{t-k+1} \quad \text{.....(5.7)}$$

where $m = 2$, and $n = 1$.

Second when using the LPM in the modeling process, the seasonal mean of flows for each date of the year (1 to 365) was obtained and the departures from the mean was calculated for both input and output variables over the calibration period. The mean values obtained for the output variable was then smoothed by the Fourier smoothing technique. The smoothing of the mean values was carried using different harmonics and seven harmonics were found to be accepted (Figure (5.16)). On the other hand, input output departure series were modeled using the TLM concept. The identified model of the input-output departure series can be shown by Equation (5.7), where $m = 2$, and $n = 1$, and the LPM can be viewed as a Modified LPM.

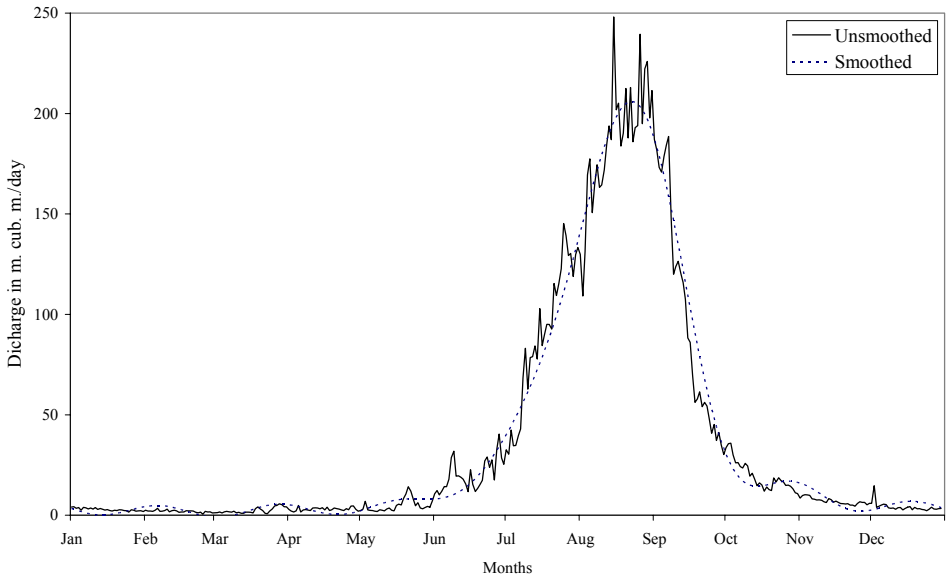


Figure (5.16): Smoothed and unsmoothed seasonal means inflow hydrographs at forecasting site no.3 (inflows to Khashm El-Girba reservoir).

Finally when using the ANNs, The optimum number of model inputs in this block was determined using the same procedure that followed in block (1). A general representation of one-step-ahead streamflow forecast model can be shown as follows:

$$q_t = g_{non}(h_{t-1}, h_{t-2}, k_t, k_{t-1}) \quad \text{.....(5.8)}$$

where $q_t, h_t, \& k_t$ are the inflows into Khashm El-Girba reservoir, the streamflow of Wed El-Heleiw station, and the streamflow of Kubur station, at time step t, respectively and g_{non} represents a nonlinear transformation to be approximated by a neural network.

The optimum network structure identified was (4, 4, 1), with one hidden layers, and it contains 4 neurons (16 weights), bias weights were included in the hidden and output layers (the total number of model parameters were 24) (Figure (5.17)).

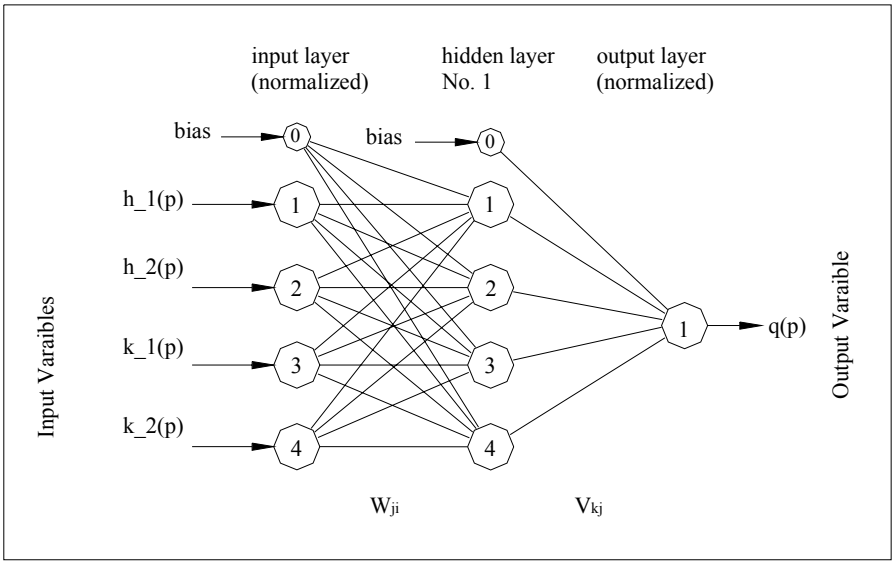


Figure (5.17): Schematic diagram shows the structure of the identified ANN model at forecasting site no.3 (inflows to Khashm El-Girba reservoir).

In the calibration process of the ANN model, initial values of weights, bias, and the factor of the Logistic function were set into random values. During calibration process and when the rate of increasing of efficiency slow down, the bias weights and the Logistic function factors were excluded from the optimization process. Two different activation functions were used; the Logistic function (Equation (3.13)) was used in the hidden layer while in the output layer the Symmetric Logistic function (Equation (3.8)) was used. The total time consumed to optimize the model parameters was about two hours. It is worth mentioning that different starting points were tried during the calibration process. Although, different values were found in all trials for each weight, all trials end up with the same efficiency.

5.3.2.1.2 *Calibration and Verification Results*:-

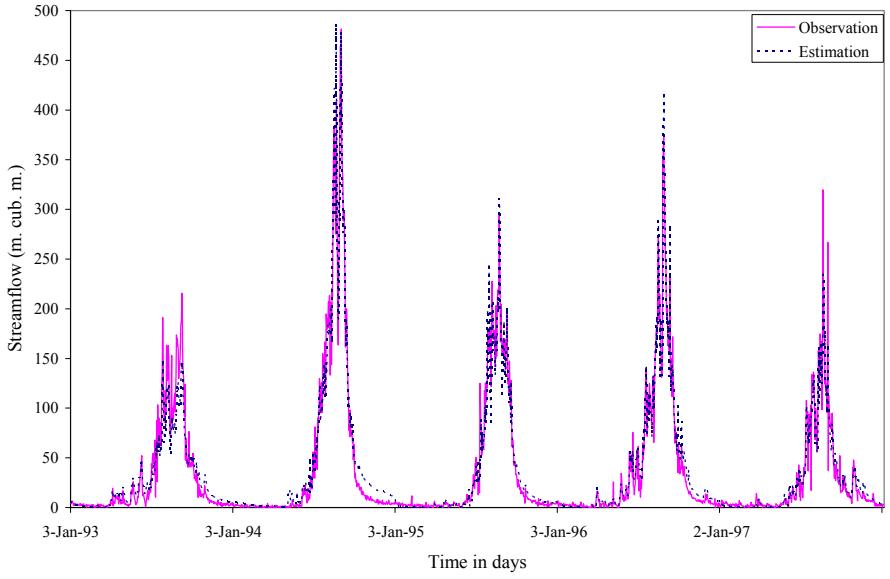
The calibration and validation period performances of all models are shown in Table (5.3), and details of their performance can be seen graphically in Figures (5.18) to (5.21).

Table (5.3a): Calibration and verification performance of the TLM, MLPM, and ANN at forecasting site no.3 (inflows to Khashm El-Girba reservoir),(Using 5 years (1993-1997) for calibration period and one year (1998) for verification period).

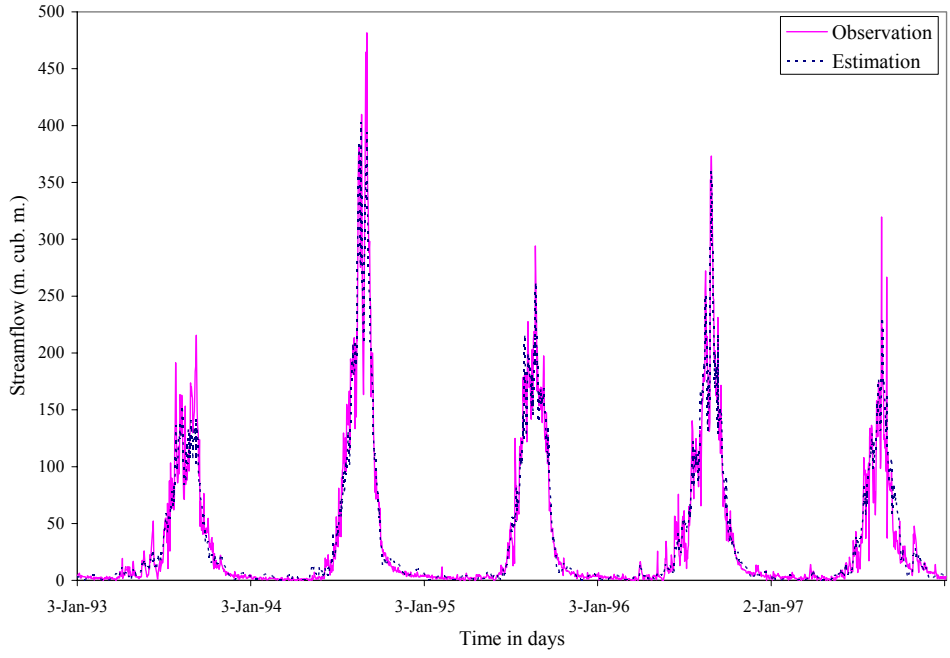
	Calibration			***Verification***		
	TLM	**MLPM**	**ANN**	**TLM**	**MLPM**	**ANN**
MIN. E	-135.25	-97.21	-120.96	-450.33	-130.57	40.12
MAX.E	127.59	132.60	92.68	60.90	166.00	-146.16
SSE(*10^3)	374.700	389.754	275.890	707.394	140.836	245.079
RMSE	14.33	14.62	12.25	44.14	19.70	25.46
R^2	0.95	0.95	0.97	0.95	0.97	0.98
Eff (%)	95.32	95.13	96.55	86.33	97.28	95.26
Bias (%)	2.55	-0.38	0.01	19.26	-3.26	-15.36
%MF	-0.59	-18.21	-10.35	60.10	1.85	-12.59

Table (5.3b): Calibration and verification performance of the TLM, MLPM, and ANN at forecasting site no.3 (inflows to Khashm El-Girba reservoir), (Using 4 years (1993-1996) for calibration period and one year (1997) for verification period).

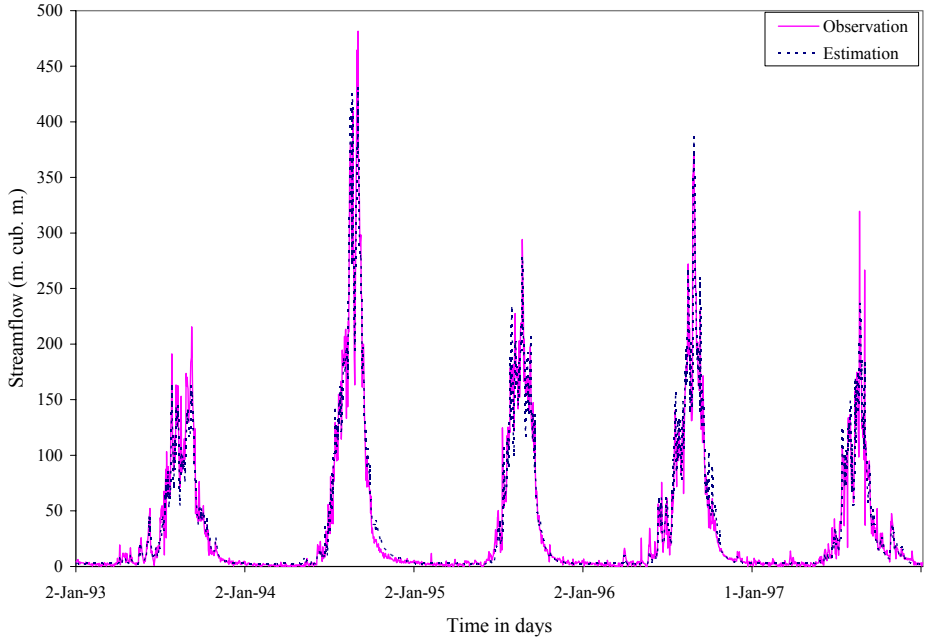
	Calibration			***Verification***		
	TLM	**MLPM**	**ANN**	**TLM**	**MLPM**	**ANN**
MIN. E	-127.08	-74.07	-136.72	-83.04	-118.52	-86.38
MAX.E	129.24	120.05	64.57	116.11	98.25	115.06
SSE(*10^3)	310.49	338.67	240.78	667.76	145.27	978.03
RMSE	14.59	15.24	12.85	13.56	20.01	8.19
R^2	0.96	0.95	0.97	0.91	0.84	0.92
Eff (%)	95.70	95.31	96.66	90.96	80.33	87.14
Bias (%)	2.80	-15.40	-6.29	0.58	12.45	18.41
%MF	-0.59	-18.21	-13.55	164.10	68.01	-9.48


Observation
Estimation
Streamflow (m. cub. m.)
500
450
400
350
300
250
200
150
100
50
0
3-Jan-93
3-Jan-94
3-Jan-95
3-Jan-96
2-Jan-97
Time in days

(5.18a)

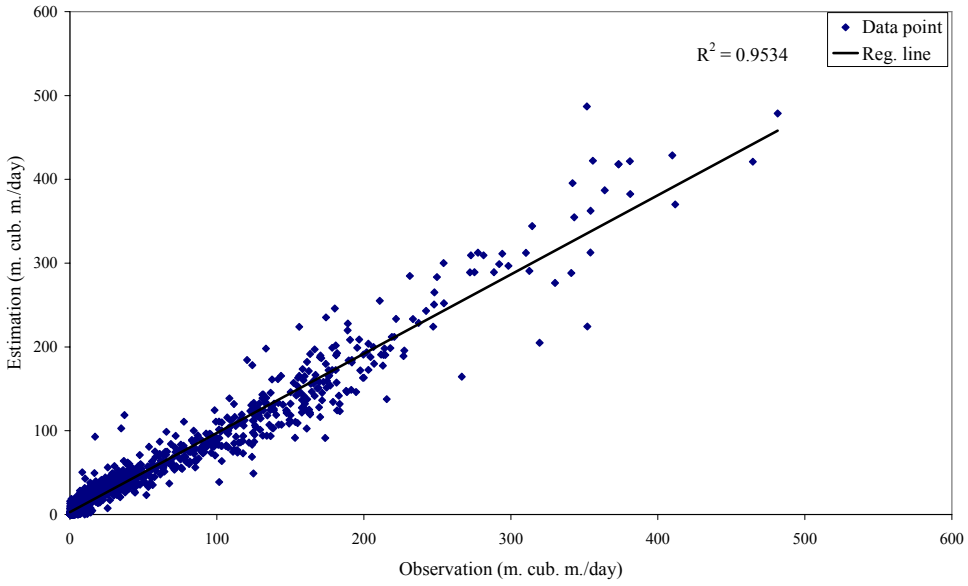

Observation
Estimation
Streamflow (m. cub. m.)
500
450
400
350
300
250
200
150
100
50
0
3-Jan-93
3-Jan-94
3-Jan-95
3-Jan-96
2-Jan-97
Time in days

(5.18b)

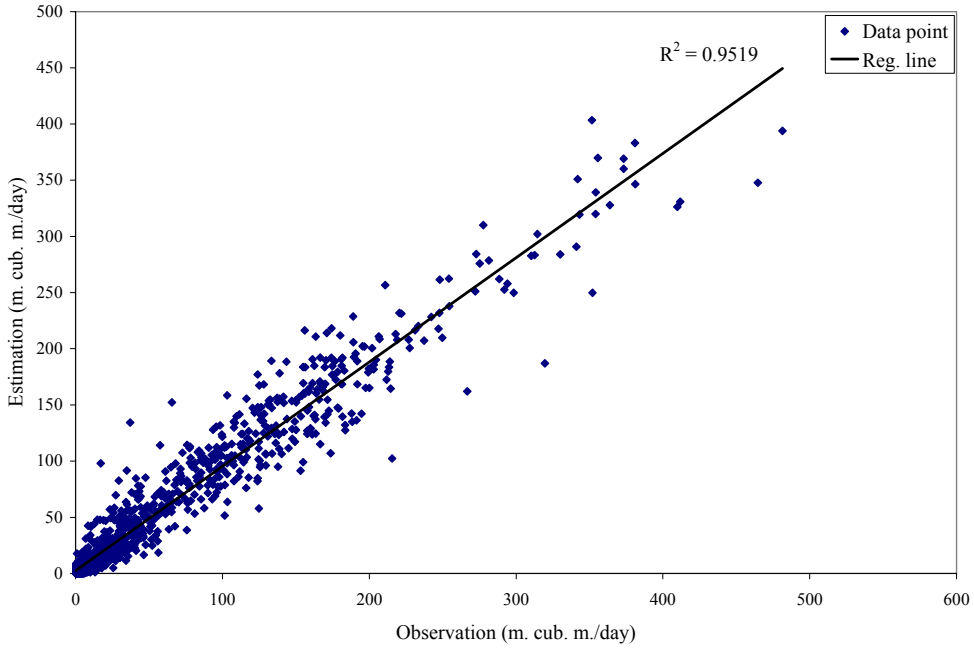


(5.18c)

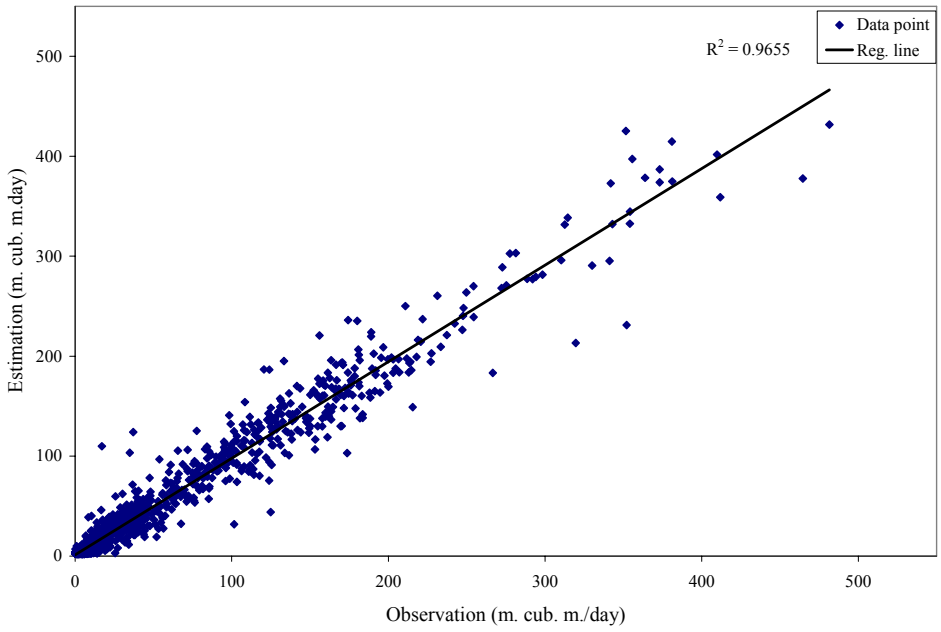
Figure (5.18): Observed and estimated daily inflow hydrographs at forecasting site no.3 (inflows to Khashm El-Girba reservoir) during calibration period using: (a) TLM, (b) MLPM, and (c) ANNs.



(5.19a)

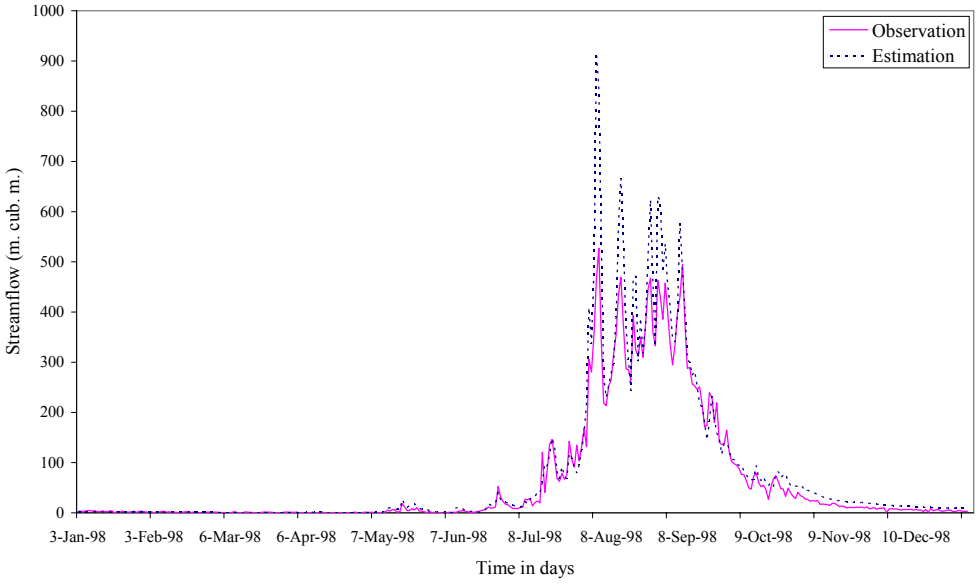


(5.19b)

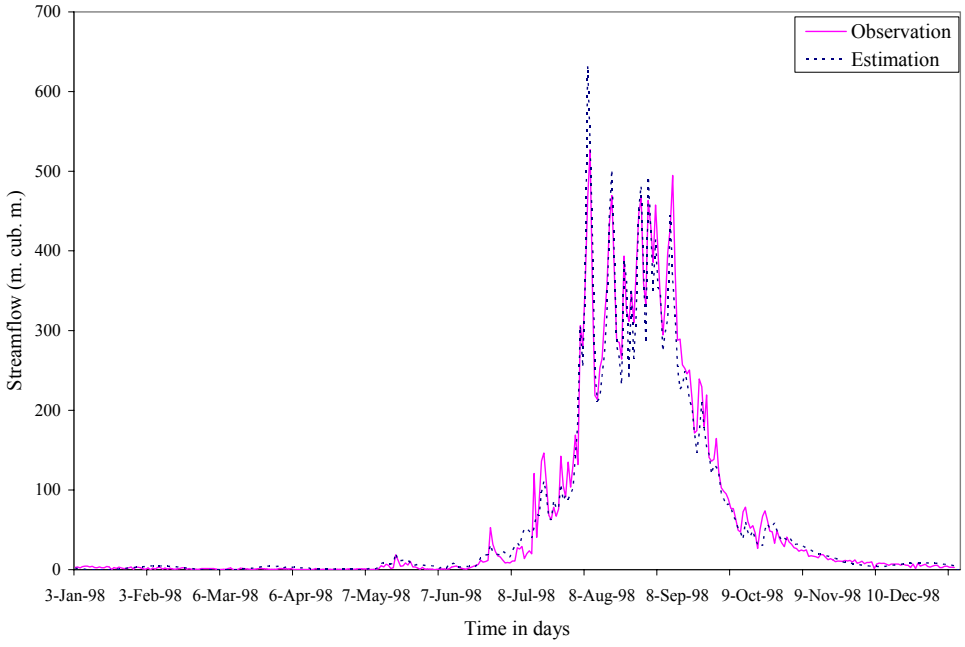


(5.19c)

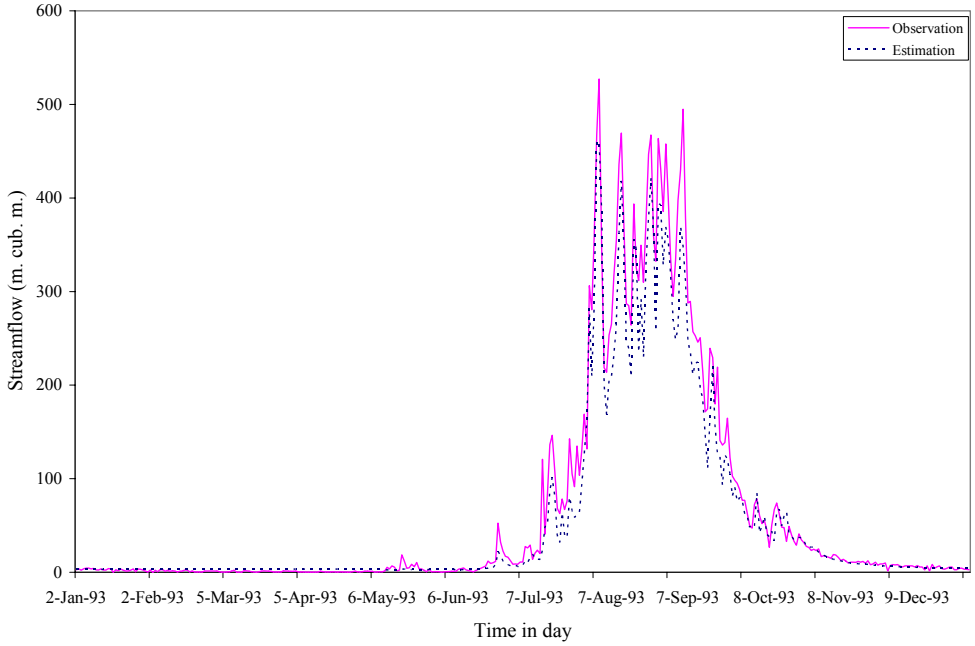
Figure (5.19): Observed and estimated daily inflow scatterplots at forecasting site no.3 (inflows to Khashm El-Girba reservoir) during calibration period using: (a) TLM, (b) MLPM, and (c) ANNs.


Observation
Estimation
Streamflow (m. cub. m.)
1000
900
800
700
600
500
400
300
200
100
0
3-Jan-98
3-Feb-98
6-Mar-98
6-Apr-98
7-May-98
7-Jun-98
8-Jul-98
8-Aug-98
8-Sep-98
9-Oct-98
9-Nov-98
10-Dec-98
Time in days

(5.20a)

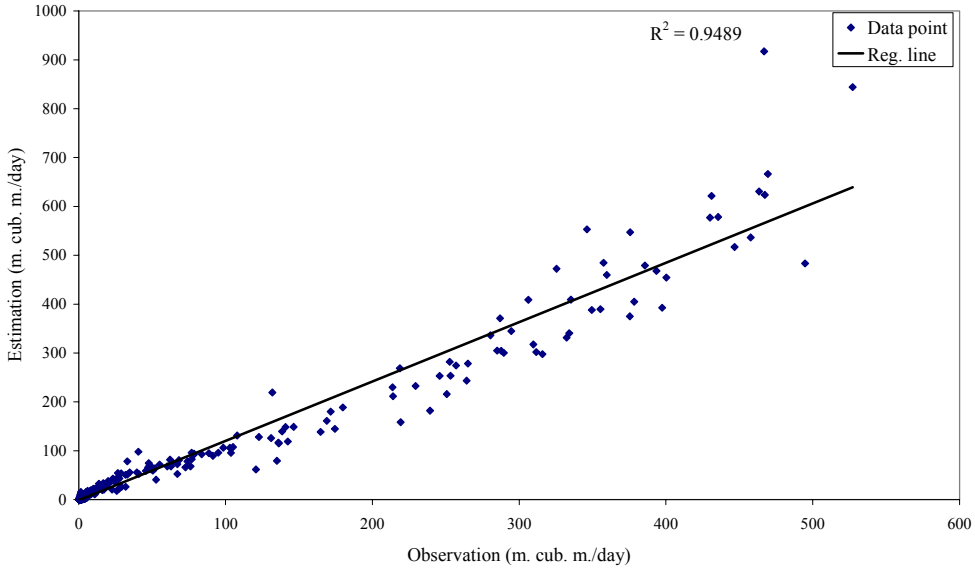

Observation
Estimation
Streamflow (m. cub. m.)
700
600
500
400
300
200
100
0
3-Jan-98
3-Feb-98
6-Mar-98
6-Apr-98
7-May-98
7-Jun-98
8-Jul-98
8-Aug-98
8-Sep-98
9-Oct-98
9-Nov-98
10-Dec-98
Time in days

(5.20b)

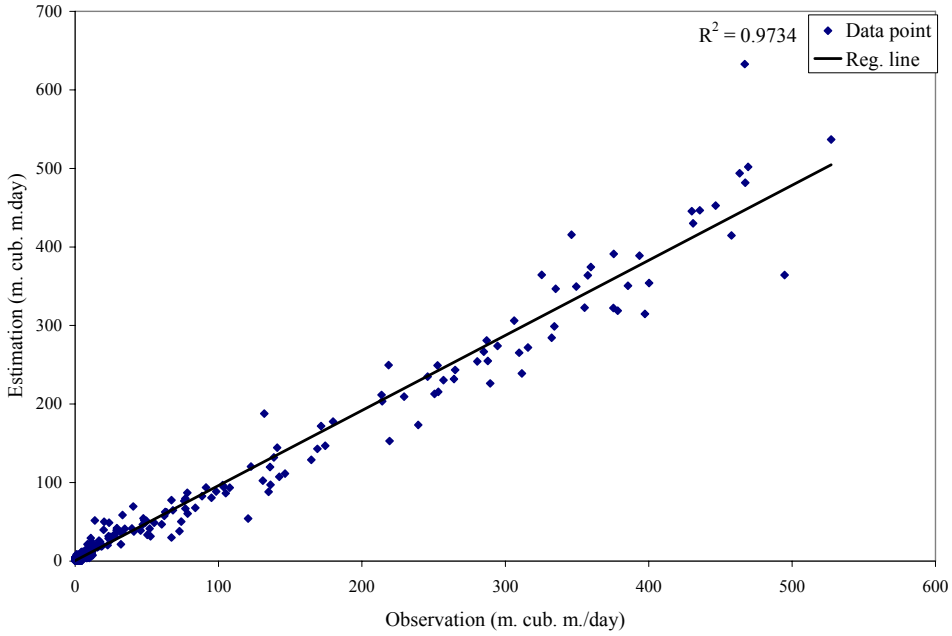


(5.20c)

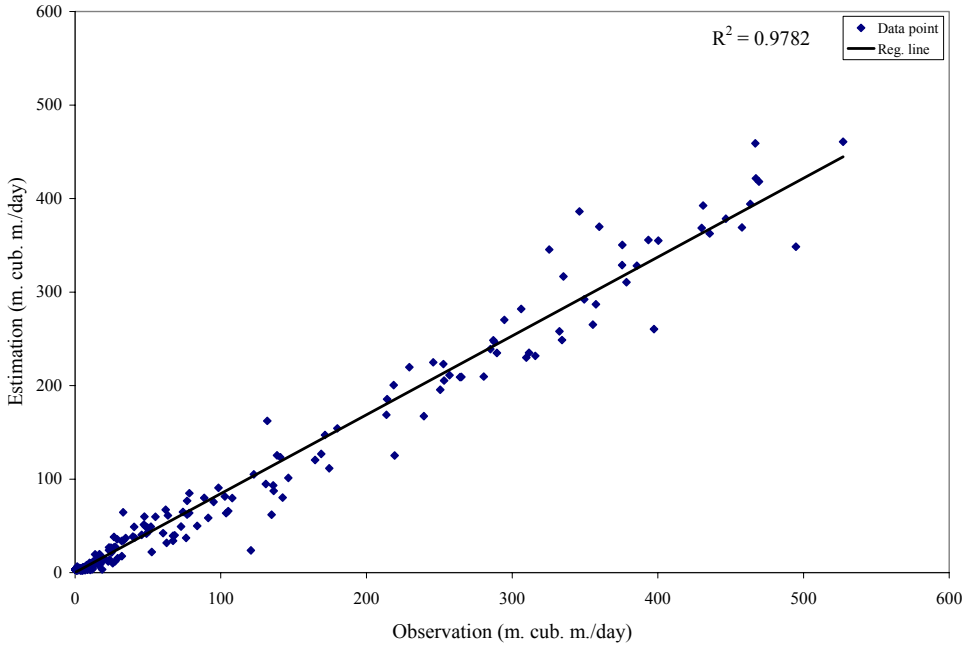
Figure (5.20): Observed and estimated daily inflow hydrographs at forecasting site no.3 (inflows to Khashm El-Girba reservoir) during verification period using: (a) TLM, (b) MLPM, and (c) ANNs



(5.21a)



(5.21b)



(5.21c)

Figure (5.21): Observed and estimated daily inflow scatterplots at forecasting site no.3 (inflows to Khashm El-Girba reservoir) during verification period using: (a) TLM, (b) MLPM, and (c) ANNs

5.3.2.2 *Forecast of Flows at Kilo_3 Station*:-

5.3.2.2.1 *Model Identification*:-

The data used to build block (4) contains daily discharges released from Khashm El-Girba reservoir (as an input variables) accompanied with the daily inflows into Kilo_3 station (as an output variable). Only the common period of data, which extend from 10 June to 31 October annually, was considered in the model, since there were no flows out of this period.

First when using the TLM, the optimum number of input lags was determined using stepwise approach, by calibrating the model each time an input lag interred. The resulting models were assessed considering their performance and the parameters stability. The identified model of the input-output departure series can be shown by Equation (5.3), where $m = 2$, and the LPM can be viewed as a Modified LPM.

Second when using the LPM in the modeling process, the seasonal mean of flows for each date of the year (1 to 365) was obtained and the departures from the mean were calculated for both input and output variables over the calibration period. The mean values obtained for the output variable were then smoothed by the Fourier smoothing technique. The smoothing of the mean values was carried using different harmonics and ten harmonics were found to be accepted (Figure (5.22)). On the other hand, input output departure series were modeled using a combination of the TLM and the AR model, since the results of this process when using the TLM show poor performance. The identified model can be shown by Equation (5.4), where $m = 1$, and $n = 1$:

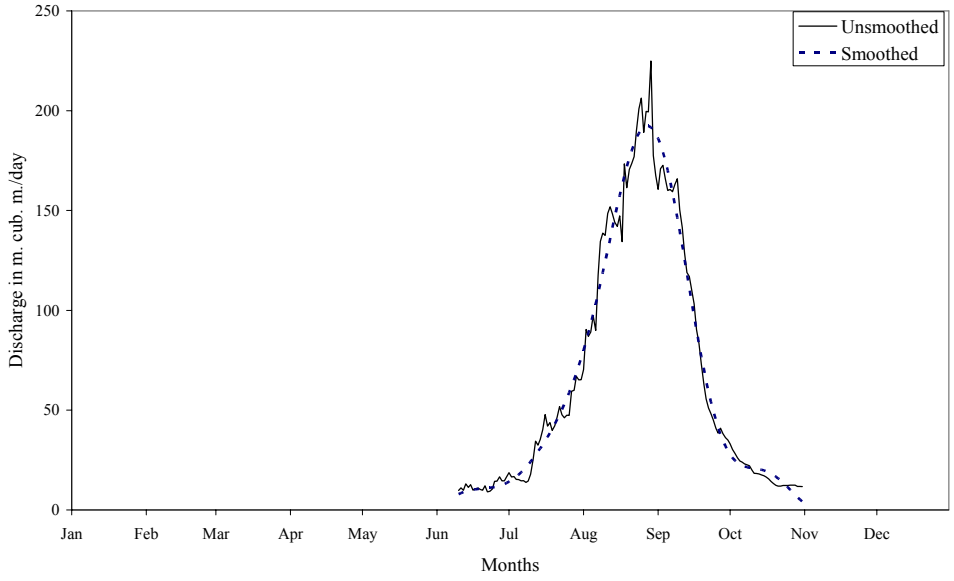


Figure (5.22): Smoothed and unsmoothed seasonal means flow hydrographs at forecasting site no.4 (Kilo_3 station).

Finally when using the ANNs model, the optimum number of model inputs in this block was determined using the same procedure that followed in block (1). A general representation of one-step-ahead streamflow forecast model can be shown as follows:

$$k_t = g_{non}(q_{t-1}, q_{t-2}, k_{t-1}, k_{t-2}) \quad \text{.....(5.9)}$$

where q_t & k_t is the discharge released from Khashm El-Girba reservoir, and the streamflow of Kilo_3 station, at time step t, respectively and g_{non} represents a nonlinear transformation to be approximated by a neural network. The optimum network structure identified was (4, 9, 1), with single hidden layer, and it contains 9 neurons (36 weights), bias weights were included in the hidden and output layers (the total number of model parameters were 49) (Figure (5.23)).

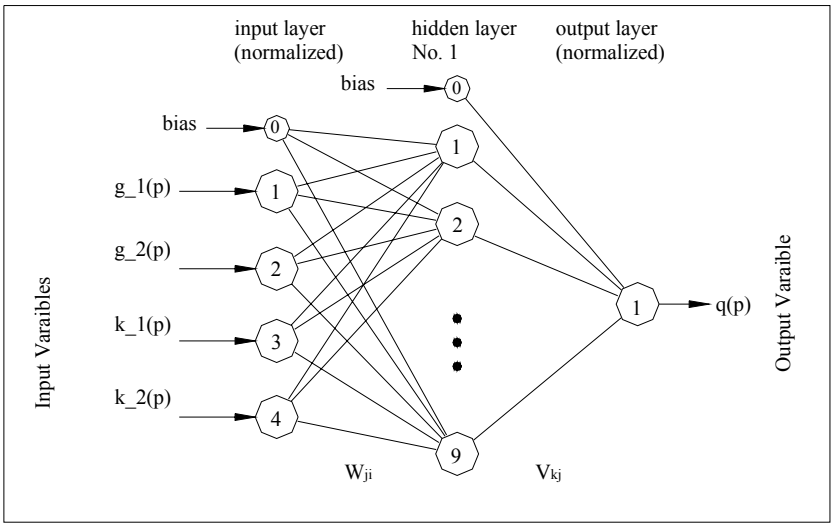


Figure (5.23): Schematic diagram shows the structure of the identified ANNs model at forecasting site no.4 (Kilo_3 Station).

In the calibration process of the ANN model, initial values of weights, bias, and the factor of the Logistic function were set into random values. During calibration process and when the rate of increasing of efficiency slows down, the bias weights and the Logistic function factors were excluded from the optimization process. Two different activation functions were used; the Symmetric Logistic function (Equation (3.13)) was used in the hidden layer while in the output layer the Logistic function (Equation (3.8)) was used. The total time consumed to optimize the model parameters was about two hours. It is worth mentioning that different starting points were tried during the calibration process. Although, different values were found in all trials for each weight, all trials end up with the same efficiency.

5.3.2.2.2 *Calibration and Verification Results*:-

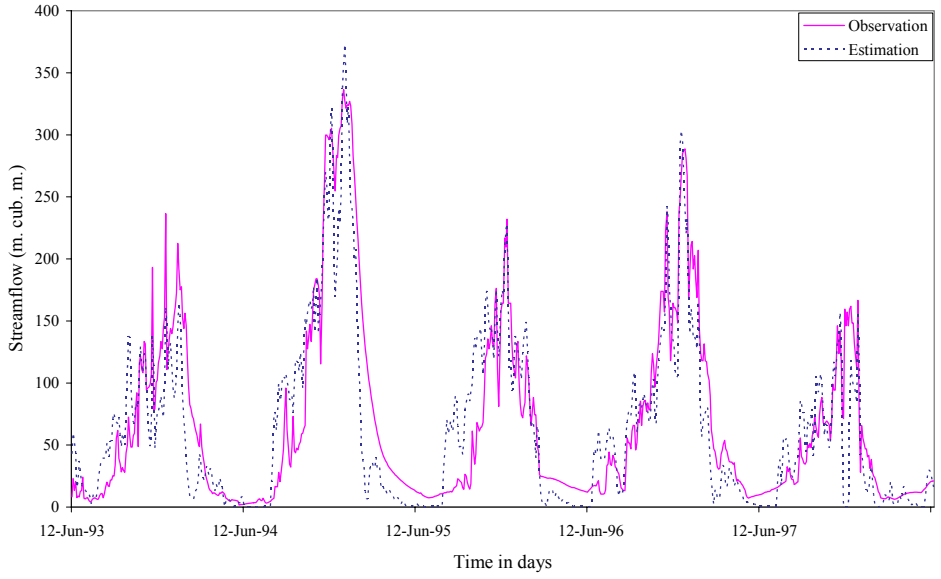
The calibration and validation periods performances of all models are shown in Table (5.4), and details of their performance can be seen graphically in Figures (5.24) to (5.27).

Table (5.4a): Calibration and verification performance of the TLM, MLPM, and ANN at forecasting site no.4 (Kilo_3 station), (Using 5 years (1993-1997) for calibration period and one year (1998) for verification period).

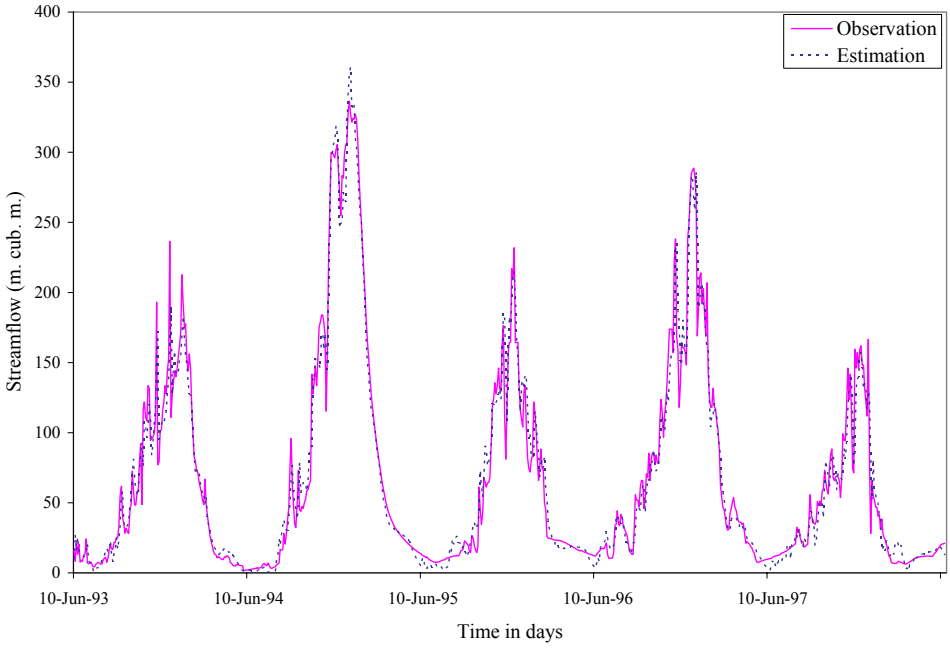
	Calibration			***Verification***		
	TLM	MLPM	ANN	TLM	MLPM	ANN
MIN. E	-101.60	-102.12	-88.08	-221.94	-37.69	-46.43
MAX.E	159.51	85.68	57.98	193.21	104.36	137.84
SSE(*10^3)	936.548	174.608	826.046	734.960	181.039	134.422
RMSE	36.12	15.58	10.73	71.94	35.71	30.77
R^2	0.77	0.95	0.98	0.77	0.98	0.96
Eff (%)	76.69	95.01	97.94	71.68	92.99	94.80
Bias (%)	-5.89	-0.28	0.01	-19.21	-14.07	6.44
%MF	-4.90	5.73	-1.89	-42.01	-13.14	-6.01

Table (5.4b): Calibration and verification performance of the TLM, MLPM, and ANN at forecasting site no.4 (Kilo_3 station), (Using 4 years (1993-1996) for calibration period and one year (1997) for verification period).

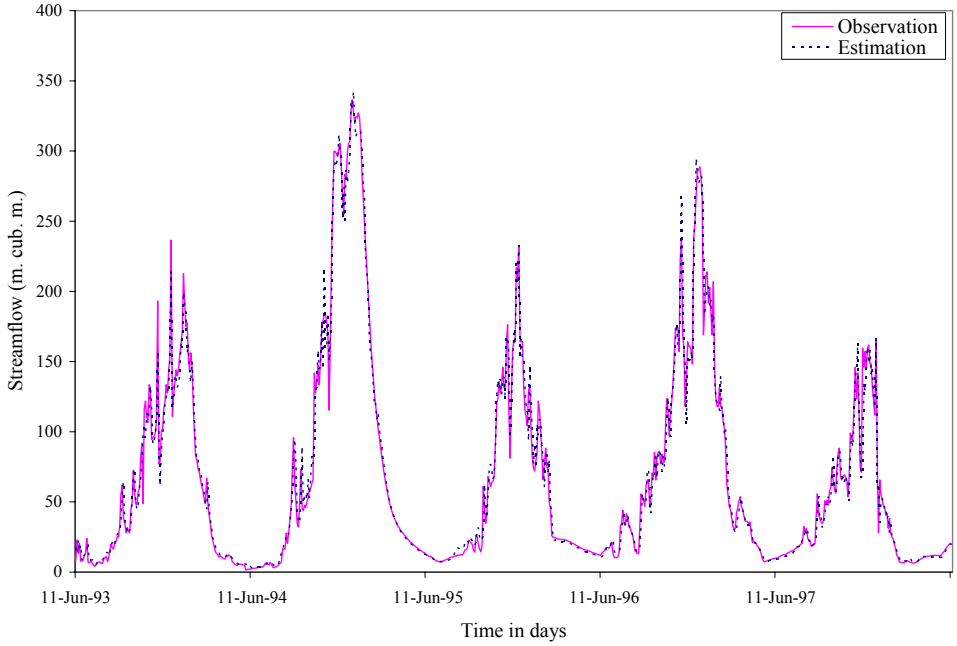
	Calibration			***Verification***		
	TLM	MLPM	ANN	TLM	MLPM	ANN
MIN. E	-102.72	-91.56	-63.87	-48.44	-52.57	-92.98
MAX.E	117.81	69.98	57.67	159.51	87.92	40.96
SSE(*10^3)	783.88	135.40	763.64	153.21	319.20	227.22
RMSE	36.92	15.35	11.52	32.85	14.94	12.65
R^2	0.79	0.96	0.98	0.491	0.88	0.87
Eff (%)	78.50	96.29	97.91	41.75	87.91	91.35
Bias (%)	-4.38	-0.03	-0.27	-12.67	-1.49	-3.73
%MF	-4.90	5.73	-1.89	67.03	150.18	-6.95


Observation
Estimation
Streamflow (m. cub. m.)
400
350
300
250
200
150
100
50
0
12-Jun-93
12-Jun-94
12-Jun-95
12-Jun-96
12-Jun-97
Time in days

(5.24a)

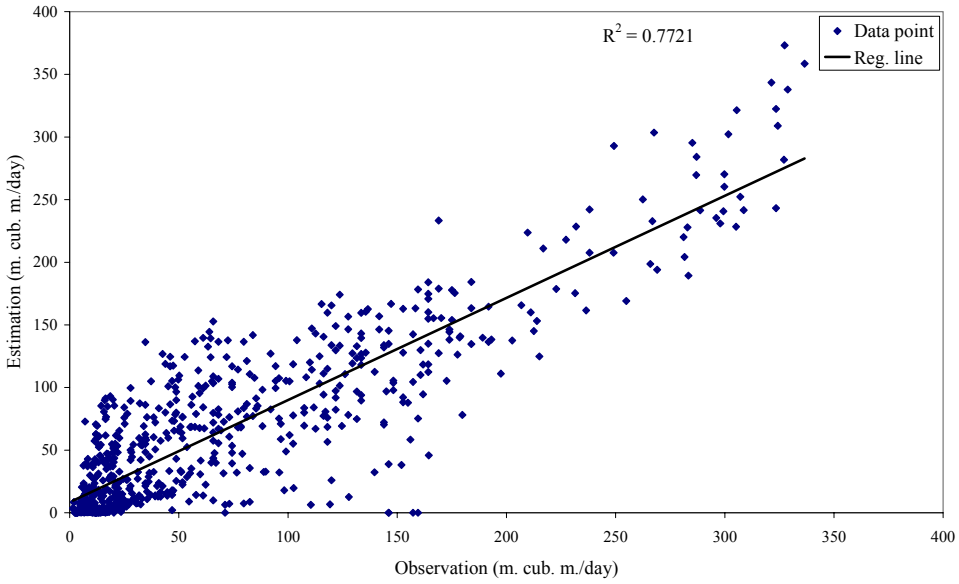

Observation
Estimation
Streamflow (m. cub. m.)
400
350
300
250
200
150
100
50
0
10-Jun-93
10-Jun-94
10-Jun-95
10-Jun-96
10-Jun-97
Time in days

(5.24b)

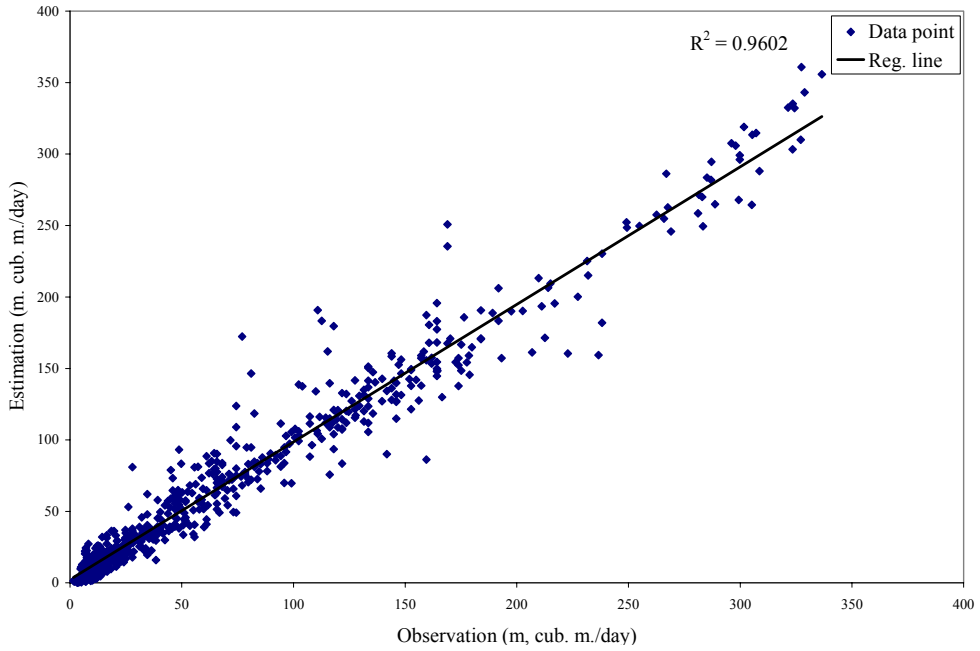


(5.24c)

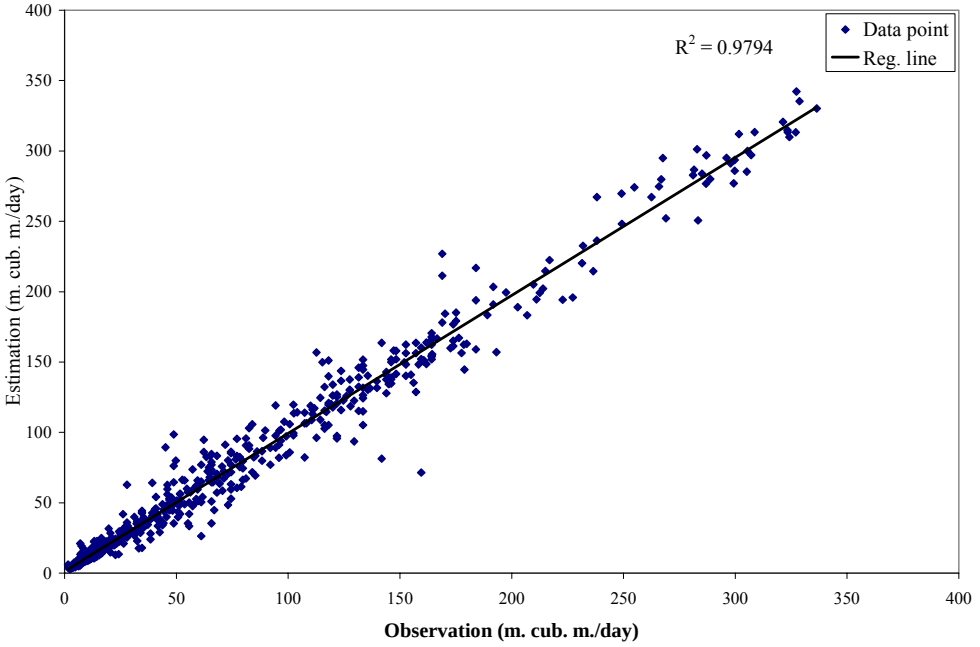
Figure (5.24): Observed and estimated daily streamflow hydrographs during calibration period at forecasting site no.4 (Kilo_3 station) using: (a) TLM, (b) MLPM, and (e) ANNs.



(5.25a)

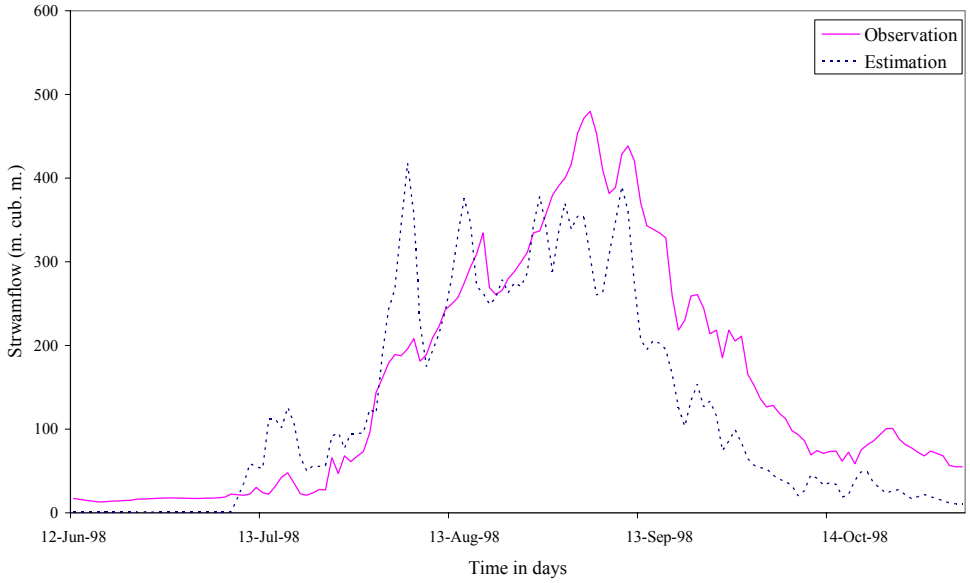


(5.25b)

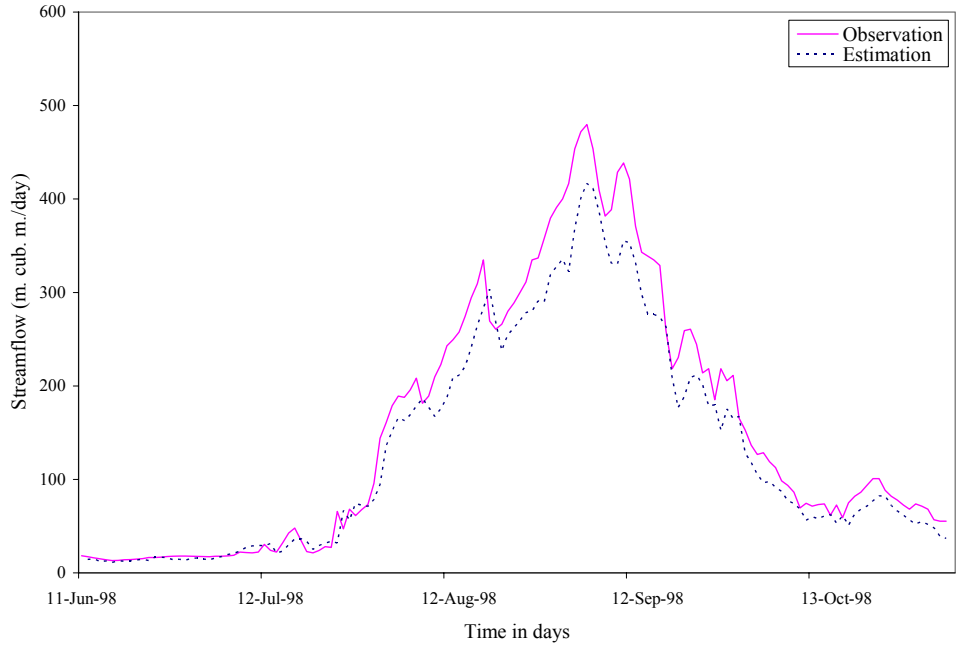


(5.25c)

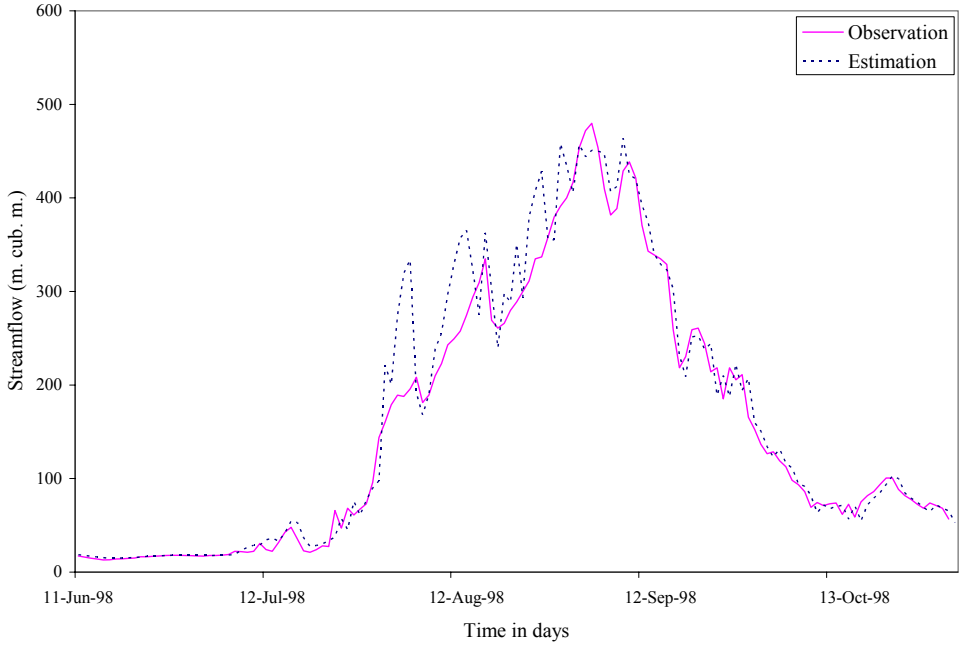
Figure (5.25): Observed and estimated daily streamflow scatterplots during calibration period at forecasting site no.4 (Kilo_3 station) using: (a) TLM, (b) MLPM, and (e) ANNs.


Observation
Estimation
Strwamflow (m. cub. m.)
600
500
400
300
200
100
0
12-Jun-98
13-Jul-98
13-Aug-98
13-Sep-98
14-Oct-98
Time in days

(5.26a)

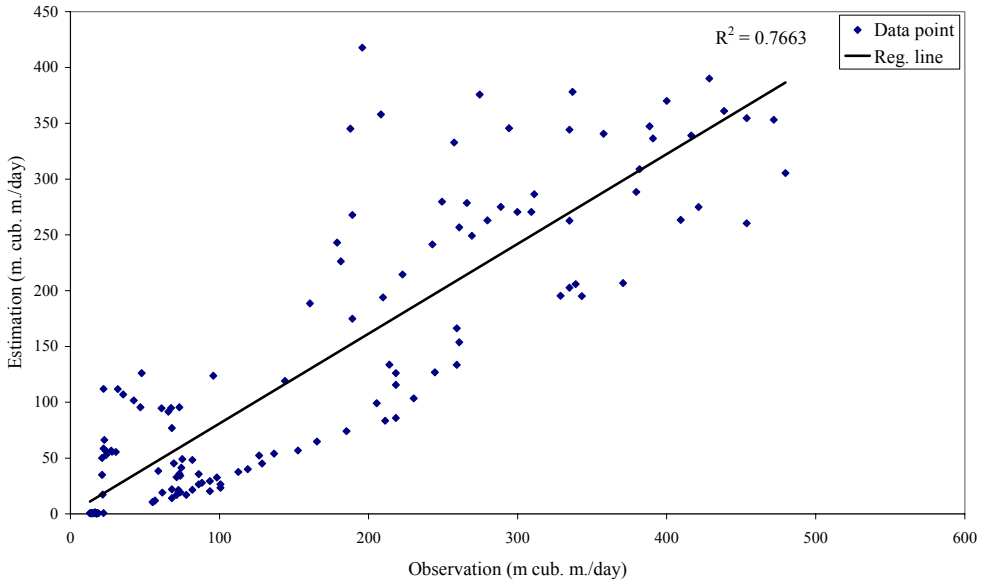

Observation
Estimation
Streamflow (m. cub. m./day)
600
500
400
300
200
100
0
11-Jun-98
12-Jul-98
12-Aug-98
12-Sep-98
13-Oct-98
Time in days

(5.26b)

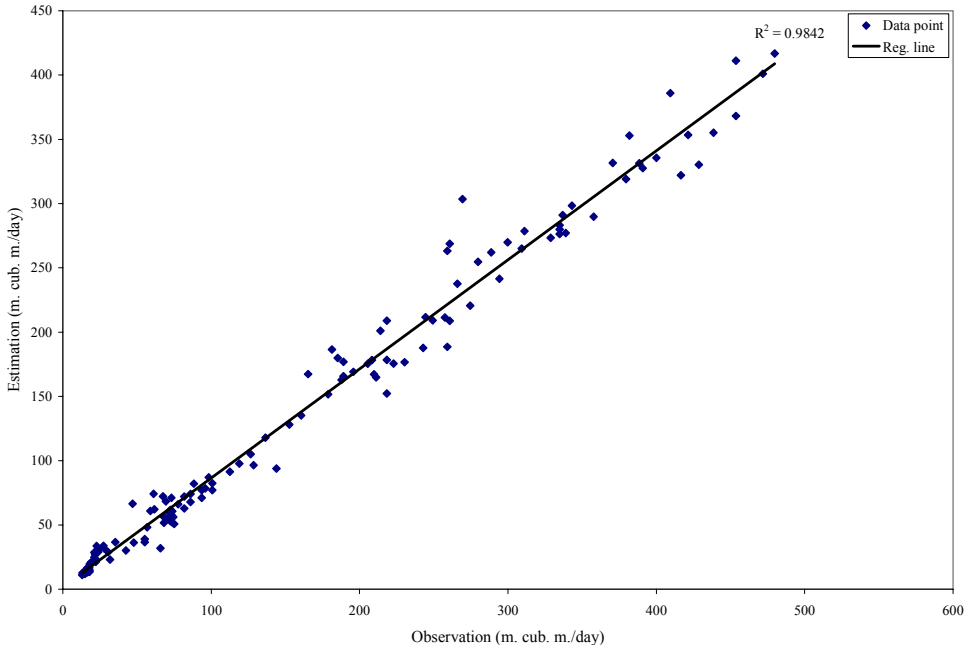


(5.26c)

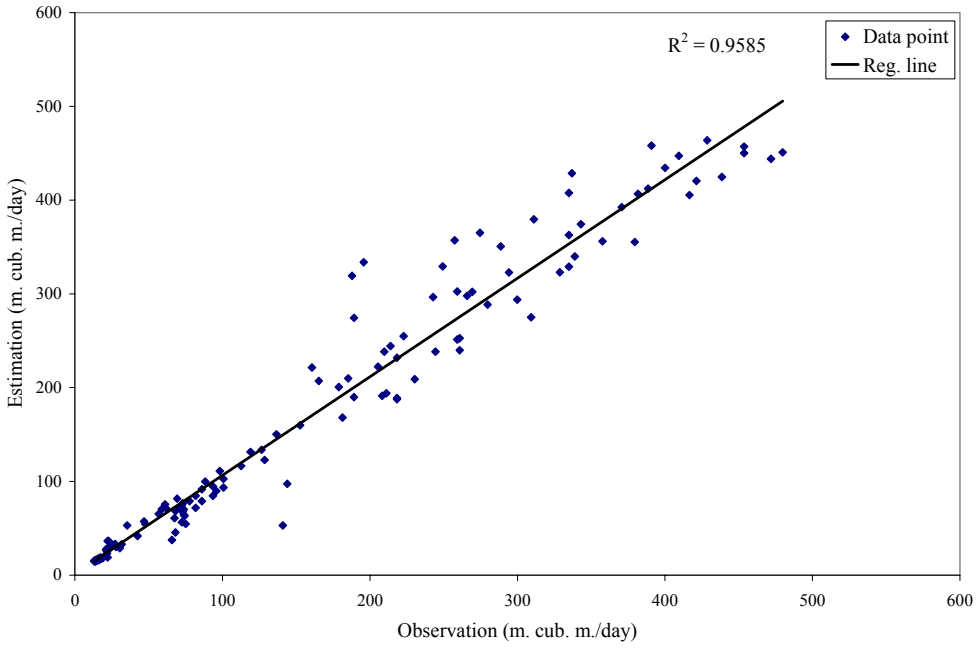
Figure (5.26): Observed and estimated daily streamflow hydrographs during verification period at forecasting site no.4 (Kilo_3 station) using: (a) TLM, (b) MLPM, and (e) ANNs.



(5.27a)



(5.27b)



(c)

Figure (5.27): Observed and estimated daily streamflow scatterplots during verification period at forecasting site no.4 (Kilo_3 station) using: (a) TLM, (b) MLPM, and (e) ANNs.

5.4 Discussion of Results:-

The statistical performance of the identified models for the five-water year calibration period and the one-water year validation period, respectively are summarized in Tables (5.1a), (5.2a), (5.3a) and (5.4a). The results are presented and discussed below.

1- The Root Mean Square Error (RMSE) statistics measure the residual variance; the optimal value is 0.0. The ANNs models tend to have the smallest RMSE during both calibration and validation at all of the forecasting sites except that at Kilo_3 station, since it has slightly worse RMSE at validation than the LPM. The MLPMs have significantly smaller RMSE than the TLM during both calibration and validation, but have worse RMSE than the ANNs models. The TLMs performance is the worst of all models during both calibration and validation. On average, the ANNs models perform best as measured by this statistics.

2- The percentage error in total volume (bias, %VE) statistics measure the percent error in volume under the observed and simulated hydrographs, summed over the data period; 0.0 is the best (+) values indicate overestimation while (-) values indicate underestimation. During calibration, all of the models perform well. During validation, the MLPM performance deteriorates just a little, the TLM performance becomes dramatically worse, while the ANNs performance stays consistently good. The ANNs performance is the best as measured by this statistics.

3- The percentage error in matching the peak flow (%MF) statistics measures the percent error in matching the maximum (peak) flow of the data record; 0.0 is the best, (+) indicate overestimation, and (-) indicate underestimation. During both calibration and validation periods, all of the models underestimate the peak flow at all of the forecasting sites, during validation, the performance deteriorates in every case, but it was just a little in the ANNs models. The worst deterioration is for the TLM (from 2.67to 42.01% underestimate); the ANNs models underestimate by approximately 22% and the MLPM

underestimate by approximately 27% as an average of all of the forecasting sites. On average, the ANNs models performance is the best as measured by this statistics.

4- The coefficient of determination (R^2) statistics measures the linear correlation between the observed and estimated flows; optimal value is 1.0. The ANNs models have the highest R^2 of all models during both calibration and validation periods except one case at Kilo_3 station. The R^2 value is the worst during validation for the TLM. The ANNs perform best as measured by this statistics.

The models carried at the forecasting site No. 3 (inflows to Khash El-girba reservoir, were selected as representative of the TLM, MLPM, and ANNs respectively Figures (5.28a) and (5.28b) present the RMSE and the bias (%VE) statistics, for each model, computed separately for each of the six water years, and presented as a function of the mean flow for that year.

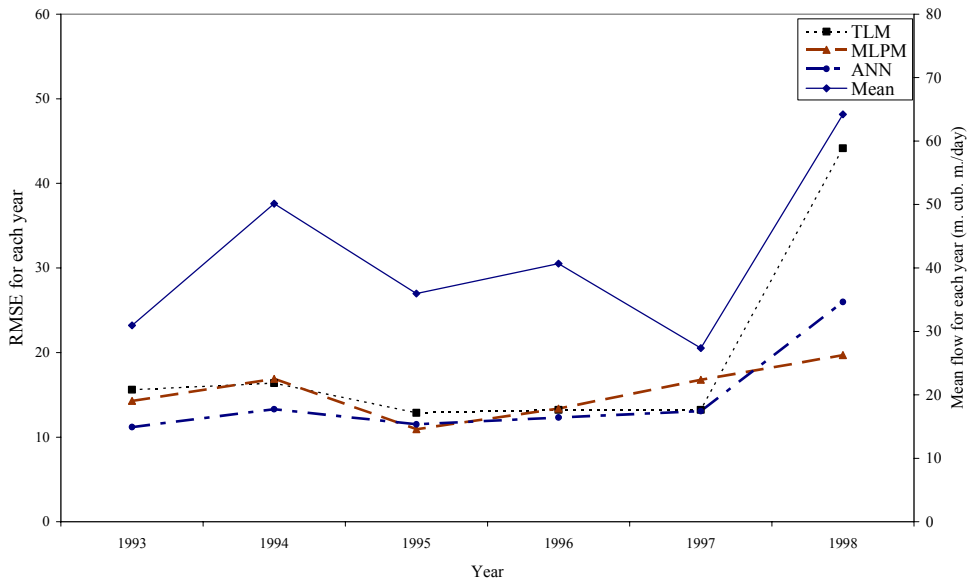


Figure (5.28a): Time series plot for the RMSE for each year for the TLM, MLPM and ANN at forecasting site no. 3 (inflows to Khashm El-Girba reservoir).

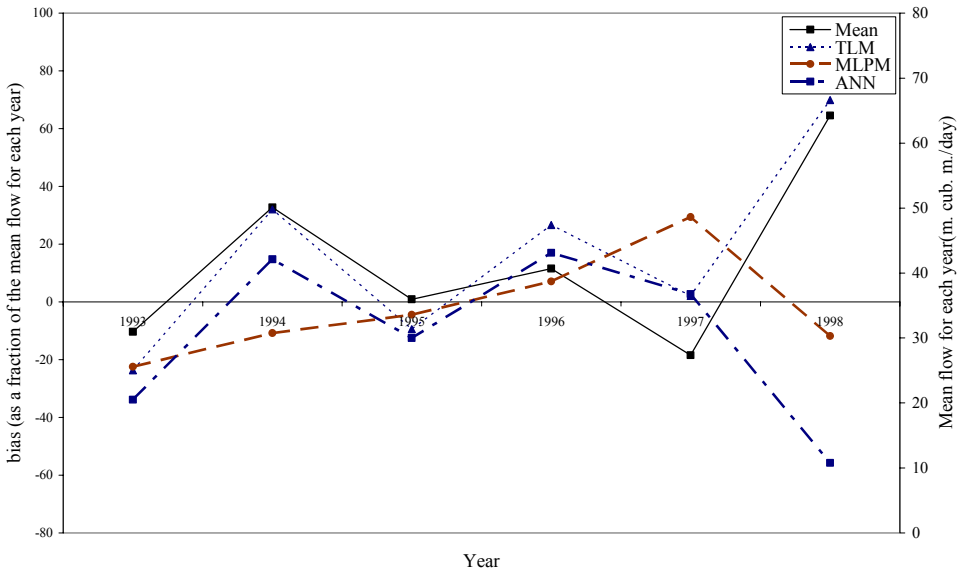


Figure (5.28b): Time series plot for the bias for each year as a fraction of the mean flow for that year for the TLM, MLPM and ANN at forecasting site no. 3 (inflows to Khashm El-Girba reservoir).

The RMSE statistics is presented in Figure (5.28a). Notice that, for all three models, there appears to be a somewhat linear correlation between the ability to match the data (residual variance measured by RMSE) and the wetness of the year; lower mean-flow years have generally lower model RMSEs. Clearly, the RMSE performance of the TLM is worse than that of the ANNs and MLPM on all years, where the TLM appears to have some trouble. This may suggest that the TLM approach has some difficulty modeling the behavior of the watershed under more extreme and variable conditions. Note also that, for the one validation year, which happens to be high "wet" year, the RMSE performance of the ANNs is worse than the MLPM. However, the RMSE performance of the ANN model on the low "wetness" calibration years is dramatically superior to that of the LPM and TLM and they follow the general trend of higher RMSE for wetter years. This may suggest that the ANN model has over-fit the data. The bias statistics is shown in Figure (5.28b). The results indicate that the ANN model has, on average, the smallest bias.

To examine these results in more detail, Figures (5.29a) to (5.29d) present the RMSE and the bias (%VE) statistics computed separately for flows above and below the six-year

mean (41.55 m. cub. m./day). This provides indication of performance on high and low flow events. Figure (5.29a) presents the RMSEA statistic (Root Mean Square Error for Above mean flows). Clearly, the TLM performs well for dry to average years and deteriorates dramatically as the "wetness" increases. Conversely, the ANN and MLPM performances are consistently good during both dry and relatively wet years. This may be a very significant finding as it may be related to the inability of a "linear" model structure to handle the more variable conditions encountered in wetter years. Figure (5.29b) presents the RMSEB statistic (Root Mean Square Error for Below mean flows); the ANN model performance is superior for all of the years except in the validation year. Figure (5.29c) presents the biasA (%VEA) statistic (present volume error for Above mean flows); the ANN and MLPM provide consistently low bias, while the TLM performance is poor. Figure (5.29d) presents the biasB(%VEB) statistic (present volume error for Below mean flows); the TLM performance is very poor. However, these results indicate that the ANN model may provide best performance under more extremes and variable conditions.

Figures (5.5), (5.12), (5.18), and (5.24) show the ability of all models used to match the hydrographs for the calibration data set (five-water years), Figures (5.7), (5.14), (5.20), and (5.26) for the validation data set. Notice the following features of these results:

1. In general the ANNs and MLPM tends to fit the higher flows quite well (on a relative error basis) during calibration and the performances deteriorate during validation as excepted, while the TLM is worse of all in the ability of fitting the high flows during both calibration and validation periods.

2. The MLPM performance is smooth during high flows, but is poor during recession and low flows. However, the inability of the simulated hydrograph to properly match the shapes of the recession and the tendency to "bottom out" during very low flows may indicate the some structural inadequacy of the model.

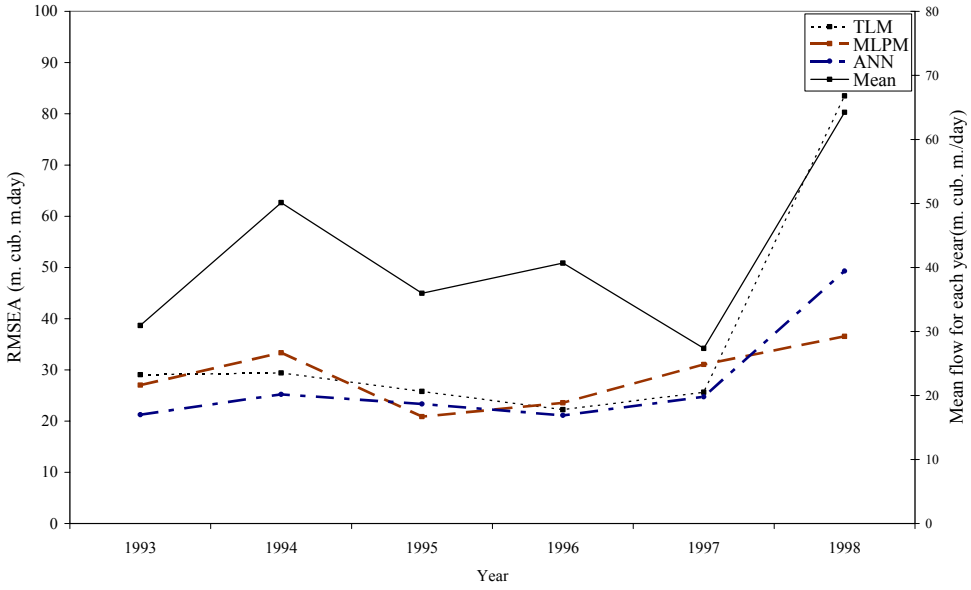


Figure (5.29a): Time series plot for the RMSEA (for above mean flows) for each year for the TLM, MLPM and ANN at forecasting site no. 3 (inflows to Khashm El-Girba reservoir).

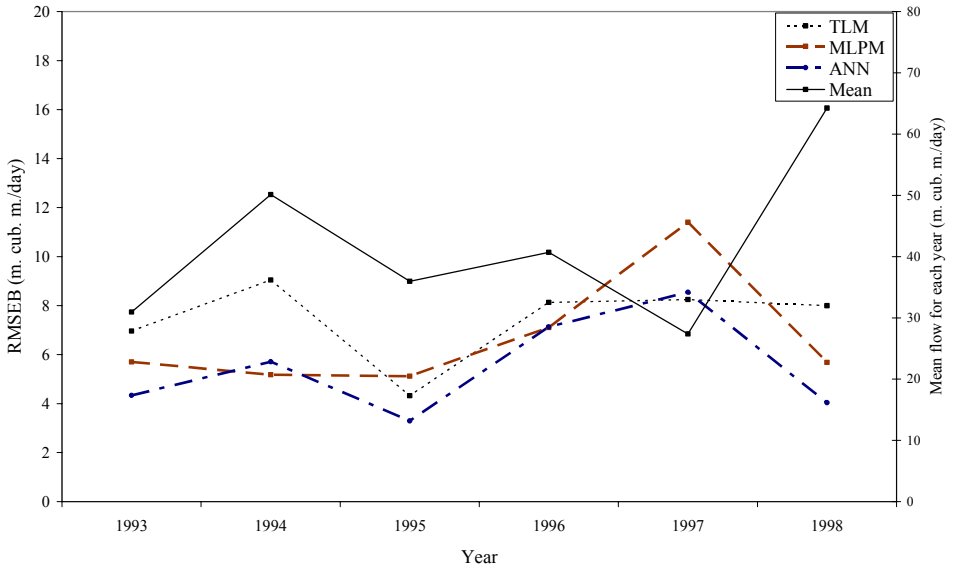


Figure (5.29b): Time series plot for the RMSEB (for below mean flows) for each year for the TLM, MLPM and ANN at forecasting site no. 3 (inflows to Khashm El-Girba reservoir).

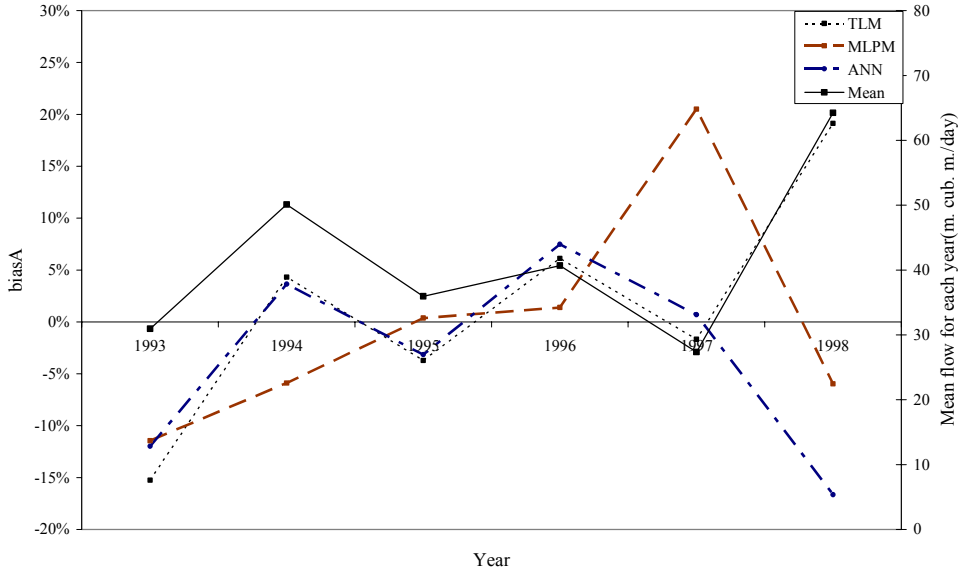


Figure (5.29c): Time series plot for the biasA (for above mean flows) for each year for the TLM, MLPM and ANN at forecasting site no. 3 (inflows to Khashm El-Girba reservoir).

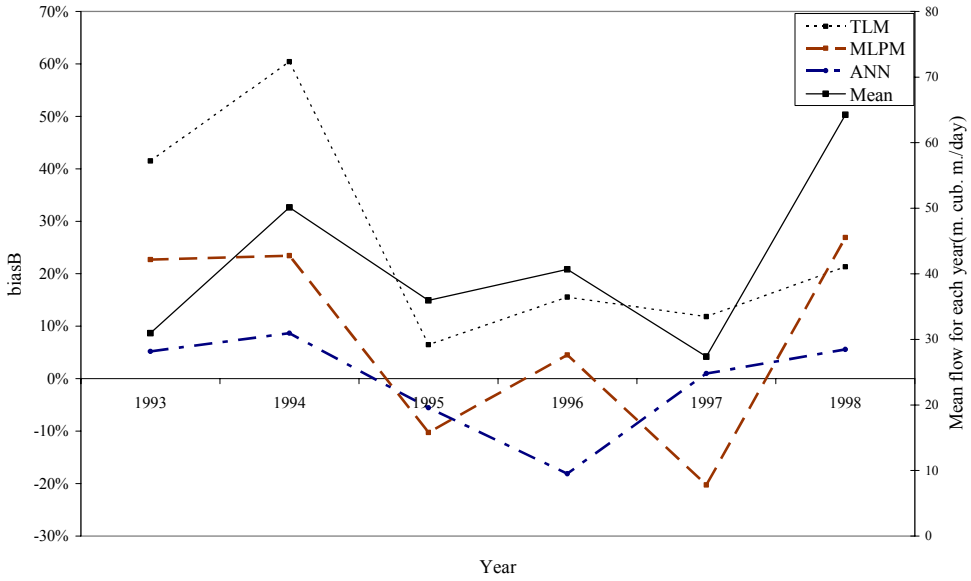


Figure (5.29d): Time series plot for the biasB (for below mean flows) for each year for the TLM, LMPM and ANN at forecasting site no. 3 (inflows to Khashm El-Girba reservoir).

3. The ANNs models match the hydrograph most closely; the high flow performance is good, while the ability to match the low and low to medium flows is better than that of the MLPM.

Estimated vs. observed flow scatter plots for all models for the calibration period are given in Figures (5.6), (5.13), (5.19), and (5.25), and Figures (5.8), (5.15), (5.21), and (5.27) for the validation period. Notice the tendency of the ANNs and LPM to underestimate in the very low flow range to overestimate in the medium flow range. The TLM tends to have the largest deviation from the 1:1 line, while the ANNs and MLPM shows the closest matching of estimated and observed over the entire flow range. As suggested before, this may indicate that the ANNs models are implicitly doing a better job of representing the nonlinearities inherent in partitioning precipitation into precipitation excess. However, there is still room for improvement.

On purpose of examining the performance of the identified models during low "wetness" or dry years, another modeling scenario was carried. Since the year 1997 shows, on average, the driest water year, in the data used in this study, it was utilized for verification period and the first 4 years from 1993 to 1996 were used for calibration period in this scenario. The statistical performances of all models after applying this scenario were summarized in Tables (5.1b), (5.2b), (5.3b) and (5.4b). Comparing the results of the two modeling scenarios (specially looking at RMSE, the bias, the %MF, R^2 and Eff), it was found that the performance of all models in all of the forecasting sites did not differ much and in some cases were almost the same during calibration period e.g. at Kilo_3 station the Eff of the ANNs were 97.94% and 97.91% for 5-year and 4-year scenarios, respectively. During verification period the performance of all models deteriorates, as expected. The worst deterioration is at Kubur station where the Eff of the MLPM deteriorates up to about 10% (from 81.46% to 72.31% for 5-year and 4-year scenarios, respectively). From this comparison, one can conclude that, on average, the ANNs models perform best even in dry "wetness" water condition.

Chapter 6

SUMMARY AND CONCLUSIONS

This study presents the potential of the most recently developed modeling approach which is known as the Artificial Neural Networks (ANNs) for the simulation of the hydrologic behavior of watersheds. An ANN is a nonlinear mathematical structure capable of representing arbitrarily complex nonlinear processes that relate the input and output (input-output mapping) of any system.

The applicability and capability of ANN in simulating both rainfall-runoff process as well as channel routing process were examined over the medium sized Atbara river sub-basin, the last major tributary of the Nile River, Sudan, using six years (1993-1998) of real daily rainfall-streamflow records. The results were then compared with other numerical simulators namely; the Total Linear Model (TLM), and the Modified Linear Perturbation Model (MLPM). A multilayer feedforward ANN model with different activation functions for each hidden layer and with quasi-Newton optimization algorithm was described as well.

The nonlinear ANN model identified appears to be superior to that of the TLM and LPM (modified) in both rainfall-runoff and channel routing processes. Although, the rainfall stations available in the region being studied are not fully representing and covering the spatial and temporal variability of the rainfall over the watershed, however, the ANN model shows satisfactory results and seems to provide a better model of the rainfall-runoff relationship than the TLM and LPM (modified). A better performance may be obtained if more stations were included or if a satellite-based remotely sensed data were used. On the other hand, all models showed satisfactory results in simulating the channel routing process.

However, the results suggest that the ANN approach may provide a superior alternative to the other numerical (traditional) simulators approach for developing input-output

simulation and forecasting models in situations that do not require to explicitly represent (modeling) the internal hydrologic structure (e.g. Evaporation, Infiltration, soil moisture etc process) of the particular watershed under investigation, and with relatively little conceptual understanding of it's hydrologic dynamics (the state of the watershed at different time steps).

Moreover, the obtained forecasts can be utilized regionally in; efficiently managing the operation of the Khashm El-Girba dam in both low and high flow periods. About two-day lead forecast can be obtained with a reasonable efficiency. One-day lead forecast was obtained in both Wed El-Heleiw and Kubur stations, and one-day lead forecast for the effect of the wave propagation of their combined hydrograph to take place at Khashm El-Girba dam. Also, the utilization of these results is possible downstream of Khasm El-Girba dam during the flood season in early warning for flood events at Atbara town (at Atbara river) since one-day lead forecast is obtained before the wave reaches Kilo_3 station, and downstream the confluence with the Nile River.

REFRENCES & BIBLIOGRAPHY

*Abrahart, R.J., See, L., Kneale, P.E., "*Using Pruning Algorithms and genetic Algorithms in a Neural Network Rainfall-Runoff Model.*", Journal of Hydroinformatics, 01.2(103-114), 1999.

Atiya, A.F., El-Shoura, S.M. Shaheen, S.I., El-Sherif, M.S., "*Comparison Between Neural Network Forecasting Techniques – Case Study: River Flow Forecasting.*", IEEE Transaction on Neural Networks, Vol. 10 No. 2 , 1999.

Attia, B., Andjelic, M., Barrett, C.B., *"Satellite Based Distributed Operational Hydrometeorological Forecasting System for the Nile River"*, on the Proceedings of the VIII International Water Resources Association Congress on Water Resources on Satisfying Future National and Global Water Demands, Vol. 1 , Cairo, Egypt, 1994.

Abdo, G.M., Liang, G.C., Bruen, M., Kackroo, R.K., and O'Connor, K.M, "*River Flow Forecasting on the Nile"*, on the Proceedings of the International Conference on Protection and Development of the Nile and other Major Rivers, Vol. 1/2 (4-7-1 to 4-7-15), Cairo, Egypt, 1992.

Abdo, G.M., "*Application of UCG Models for River Flow Forecasting on the Nile"*, presented in the International Workshop on River Flow Forecasting, China, 1992.

Amarasekera, K.N., Lee, R.F., Williams, E.R., and Eltahir, E.A.B., "ENSO *and Natural Variability in the Flow of Tropical Rivers.*", Journal of Hydrology, 200 (1997) pages(24-39), 1997.

Andah, K., Siccardi, F., *"Prediction of Hydrometeorological Extremes in the Sudanese Nile Region: A need for International Co-operation"*, on the Proceeding of the International Symposium during the XXth General Assembly of the International Union of Geodesy and Geophysics on Hydrology for Water Management of Large River Basins, IAHS, No. 201, Vienna, August 11-24 1991.

Barron, M., http://www.mbarron.net/Nile/fctfl_nf.html , 1998.

Bourget, P., *"Flood Plain Management and the U.S. Army Corps of Engineering"*, http://www.nwd-wc.usace.army.mil/PB/Mozambique/Washington/R_Floodplain.ppt , department of research and development-Headquarter-U.S.Army Corps of Eng. 2002.

Chow, V.T., Maidment, D.R., Mays, L.W., *"Applied Hydrology"*, McGraw Hill, ISBN 0-07-010810-2, 1998.

Davar, K.S., Hederson, J.M., Burrell, B.C., *"Flood Damage Reduction"*, Water International, Vol. 26 No. 2, (162-176), 2001.

*Dibike, Y.B., Solomatine, D.P., "*River Flow Forecasting using Artificial Neural Networks* .",European Geophysical Society Journal of Physics and Chemistry of the Earth, MS-No. EGS1.1-99002, 1999.

Draper, N.R., Smith, H., *"Applied Regression Analysis"*, Second Edition, John Wiley & Sons, ISBN 0-471-02995-5, 1981.

Eltahir, A. B. E., *"Drought frequency analysis, of annul rainfall series, Central and Western Sudan"*, unpublished, M.Sc. thesis submitted to the Department of Engineering and Hydrology, University College Galway, Ireland, 1987.

Georgakakose, A., *"Science-Based Management System could lessen water disputes on Nile river and worldwide"*, Georgia Institute of Technology research news, http://gtreasearchnews.gatech.edu , 1996.

Gill, P.E., Murray, W., Wright, M.H., *"Practical Optimization"*, Academic Press Inc., Harourt Brace Jovanovich Publishers, ISBN 0-12-283950-1, 1981.

*Hsu, K.L., Gupta, H.V., Sorooshian,S., "*Streamflow forecasting using Artificial Neural Networks*", in the Proceedings of the American Society of Civil Engineers (ASCE) International Conference on Water Resources Engineering, Memphis, Tennessee, 1998.

*Hsu, K.L., Sorooshian, S., Gao, X., Gupta, H.V., "*Rainfall Estimation using Artificial Neural Network*", in the Proceeding of the American Metrological Society (AMS) First Conference on Artificial Intelligence, pages (28-32), Phoenix, AZ, Boston, MA. 1998.

*Hsu, K.L., Gupta, H.V., Sorooshian, S., "*Application of a Recurrent Neural Network to Rainfall-Runoff Modeling*", in the Proceedings of the ASCE International Conference on Aesthetics in Constructed environment pages (68-73), New York, 1997

*Hsu, K., Gao, X., Sorooshian, S. and Gupta, H.V., "*Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks*", Journal of Applied Meteorology, Vol. 36, No.9 pages (1176-1190), 1997.

*Hsu, K., Gupta, H.V., Gao, X., and Sorooshian, S., "*Estimation of Physical Variables from Multi-channel Remotely Sensed Imagery using a Neural Network: Application to Rainfall Estimation*.", Water

Resources Research, Vol.35, No.5 , pages(1605-1618), 1999.

Karyabwite, D.R., "*Water sharing in the Nile River Valley"*, UNEP report, Project GNV011: Using GIS/Remote Sensing for the Sustainable Use of Natural Resources, 2000.

Kottegoda, N.T., "*Stochastic Water Resources Technology"*, the McMillan Press, 1980.

Varshney, R.S., *"Engineering Hydrology"*., Second Edition, Roorkee Press, Roorkee U. P. 247667, India, 1979.

Maier, H.R., Dany, G.C., "*Neural Networks for the Prediction and Forecasting of Water Resources Variables: A review of Modeling Issues and Applications*.", Journal of Hydrology, 15(2000)pages (101-124), 2000.

Nile Water Department, "*The Nile Basin"*, Ministry of Irrigation and Water Resources publications, on the Hydrological Year Book, Sudan, 1996.

Nile Water Department, *"Flood Early Warning System – Final Report."*, Ministry of Irrigation and Water Resources–Sudan & IHE publications, Sudan, 1992.

Ogink, H.J.M., Thabet, R.A.H.A., Nur, M.E.M., *"Flow Forecasting for Flood Control and Water Management"*, on the proceedings of the VIII international Water Resources Association Congress on Water Resources on Satisfying Future National and Global Water Demands, Vol. 2 , Cairo, Egypt, November 21-25, 1994.

Rodda, J.C., Downing, R.A., Law, F.M., *"Systematic Hydrology"*, Butter worth Publishers, ISBN 0408002344, 1976.

Roskar, J., *"Assessing the Nile Water Resources Potential of the Nile River Based on Data available at the Nile Forecasting Center in Cairo"*, Gergrafski zbornik, XXXX(2000), UDC: 556.53(282.263.1) COBISS:1.01, Slovenia, 2000.

Singh, V.P., *"Computer Models of Watershed Hydrology."*, Water Resources Publications, ISBN 0-918334-91-8, 1995.

*See, L., Corne, S., Dougherty, M., and Openshaw, S., "*Some Initial Experiments with Neural Network Models of Flood Forecasting on the River Ouse*", on the Proceedings of the Second Annual Conference of Geocomputational '97&SIRC'97, University of Otago, New Zealand, 1997.

*Zealand, C.M., Burn, D.H., Simonovic, S.P., "Short Term Strreamflow Forecasting *Using Artificial Neural Networks"*, Journal of Hydrology, 214(1999), pages (32-48), 1999.

Zurada, J.M., *"Introduction to Artificial Neural Systems"*, PWS Publishing Company, an International Thomson Publishing Company, ISBN 0-534-95460-X, 1995.

APPENDICES

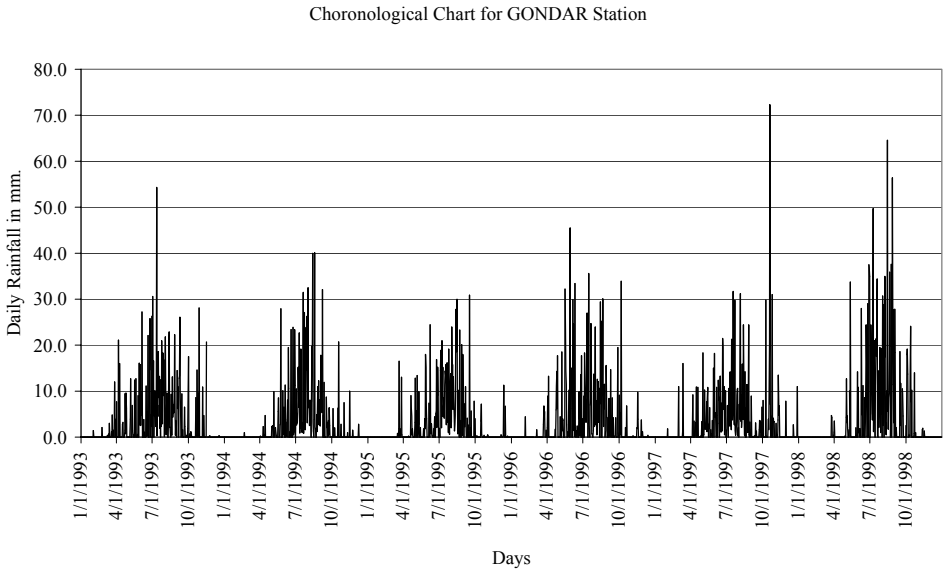
APPENDIX A (4.1)

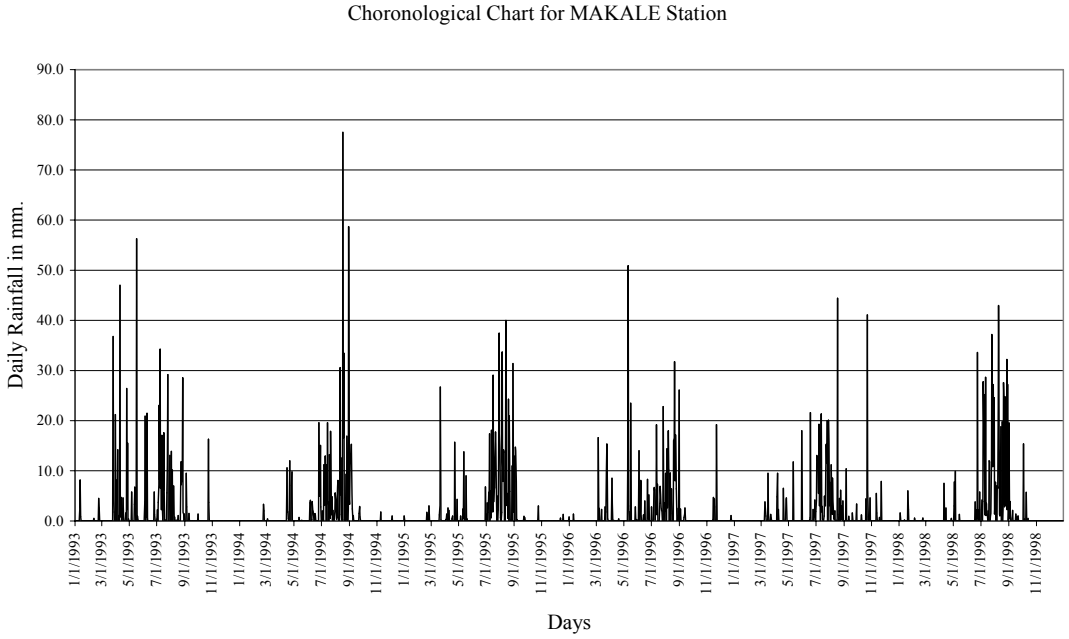
SOME STATISTICAL PROPERTIES FOR THE RAINFALL STATIONS DATA

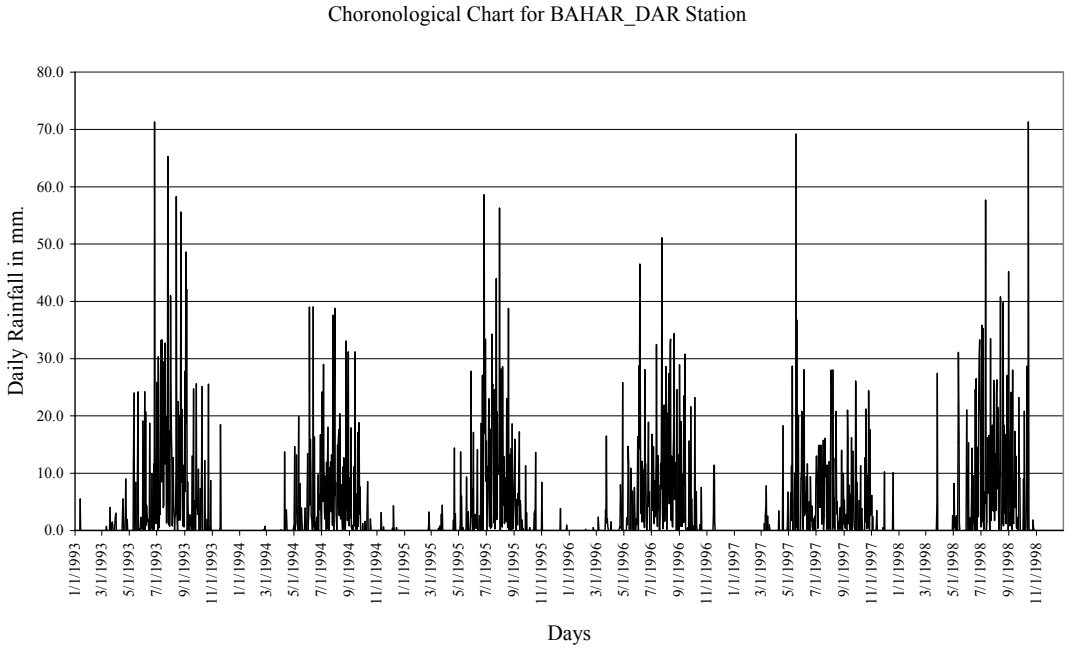
Some statistical properties for the rainfall stations data are shown in the following table:

Station	GONDAR	MAKALE	BAHAR_DA	COMBOLCH	DEBRE_MA	ADDIS_A BABA	JIMMA
N	2188	2191	2187	2191	2148	2065	2190
Range	72	78	71	68	87	65	69
Minimum	0	0	0	0	0	0	0
Maximum	72	78	71	68	87	65	69
Sum	6916	3934	7912	7089	8390	6577	9507
Mean	3.16	1.80	3.62	3.24	3.91	3.19	4.34
Std. Error of Mean	0.15	0.12	0.18	0.16	0.16	0.15	0.17
Std. Deviation	6.98	5.73	8.20	7.68	7.35	6.96	8.18
Variance	48.690	32.835	67.262	58.951	54.020	48.492	66.843
Skewness	3.445	5.274	3.498	3.509	3.158	3.244	2.932
Std. Error of Skewness	0.052	0.052	0.052	0.052	0.053	0.054	0.052

Following are time series plot for the rainfall measurements at the stations considered during the study.

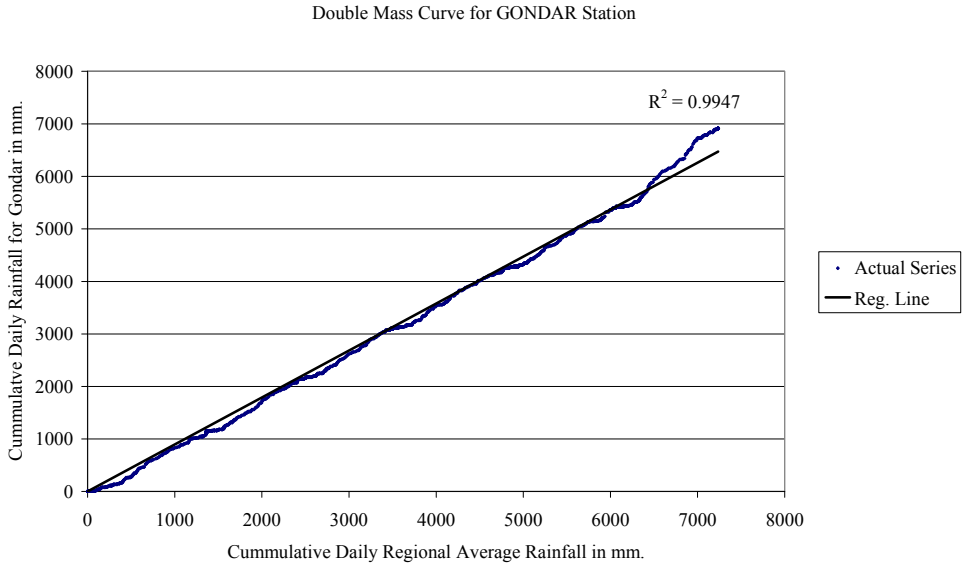


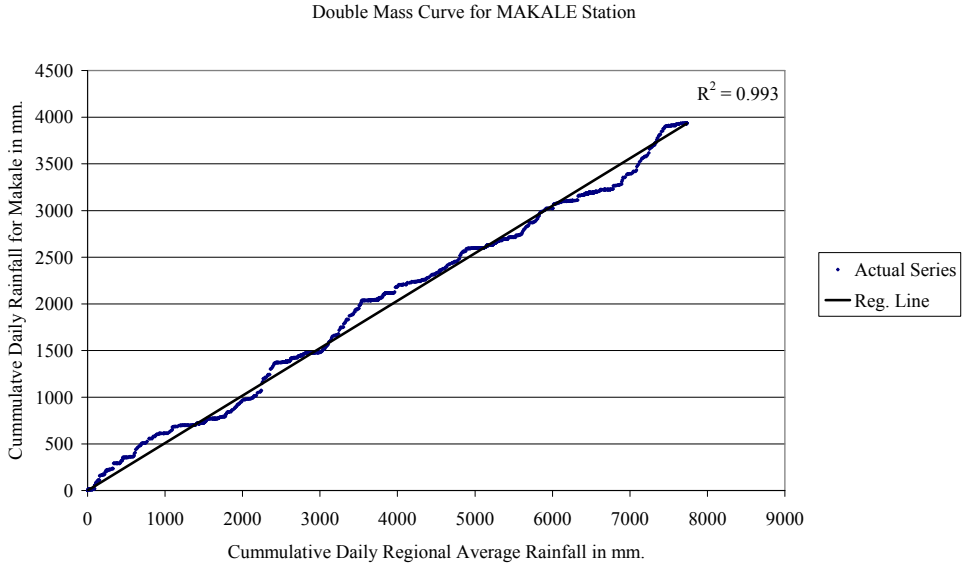


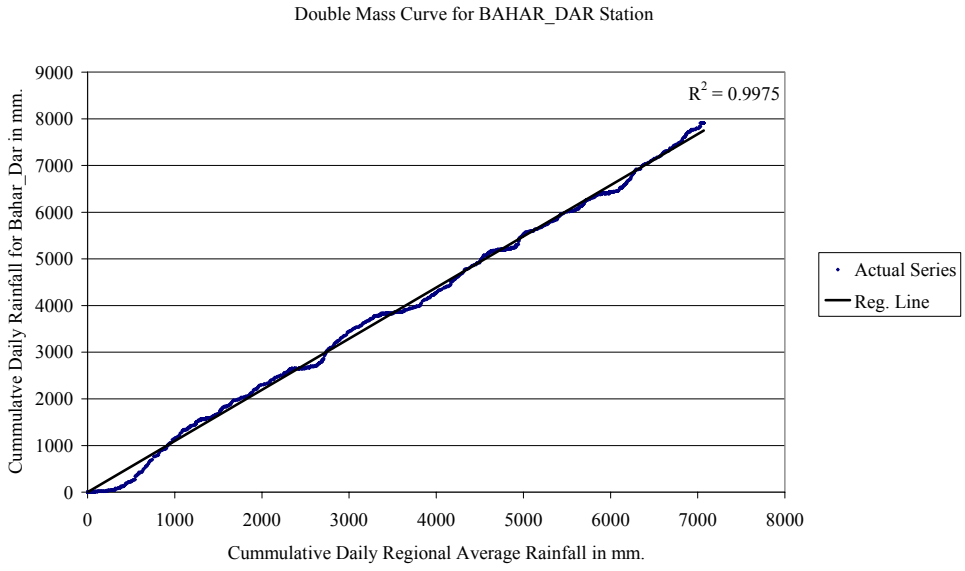

Choronological Chart for BAHAR_DAR Station
Daily Rainfall in mm.
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
1/1/1993
3/1/1993
5/1/1993
7/1/1993
9/1/1993
11/1/1993
1/1/1994
3/1/1994
5/1/1994
7/1/1994
9/1/1994
11/1/1994
1/1/1995
3/1/1995
5/1/1995
7/1/1995
9/1/1995
11/1/1995
1/1/1996
3/1/1996
5/1/1996
7/1/1996
9/1/1996
11/1/1996
1/1/1997
3/1/1997
5/1/1997
7/1/1997
9/1/1997
11/1/1997
1/1/1998
3/1/1998
5/1/1998
7/1/1998
9/1/1998
11/1/1998
Days

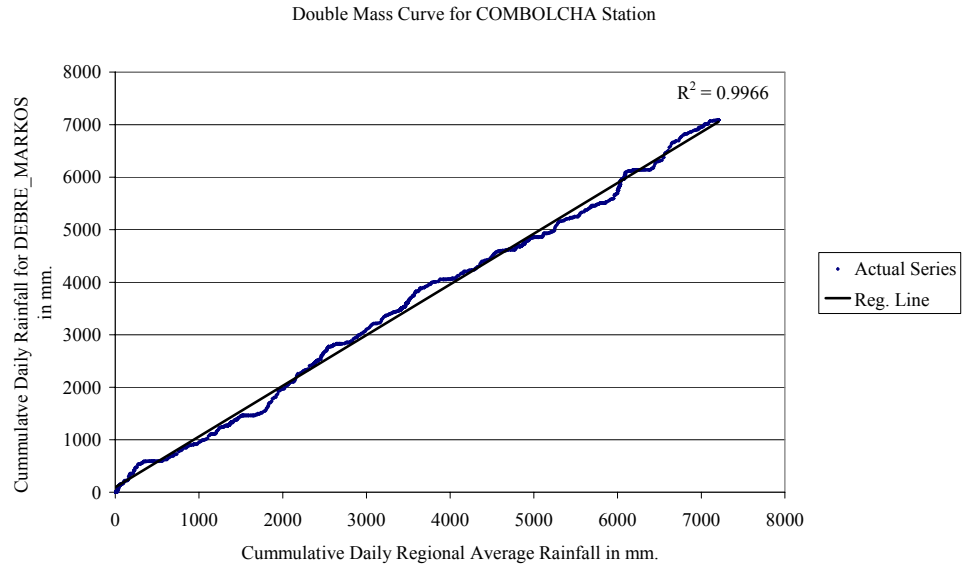
APPENDIX A (4.2)

RESULTS OF DOUBLE MASS ANALYSIS

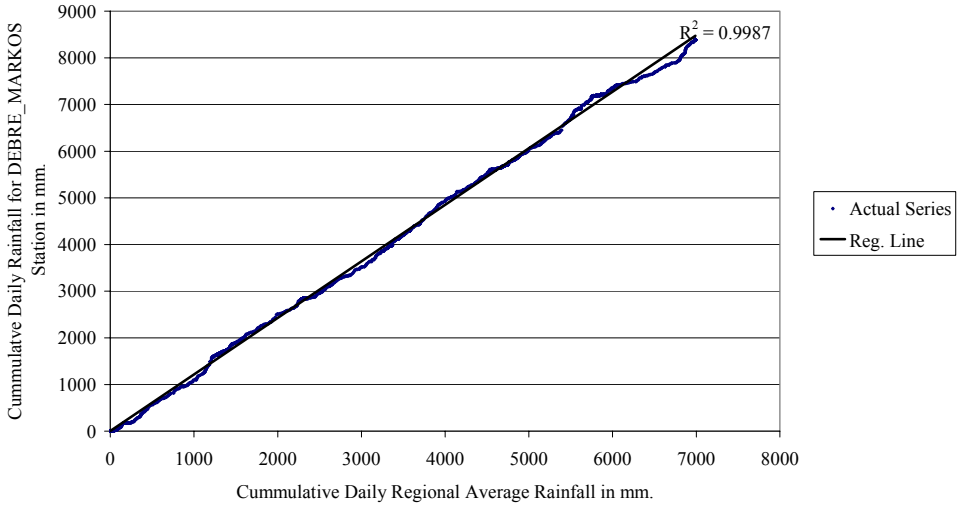




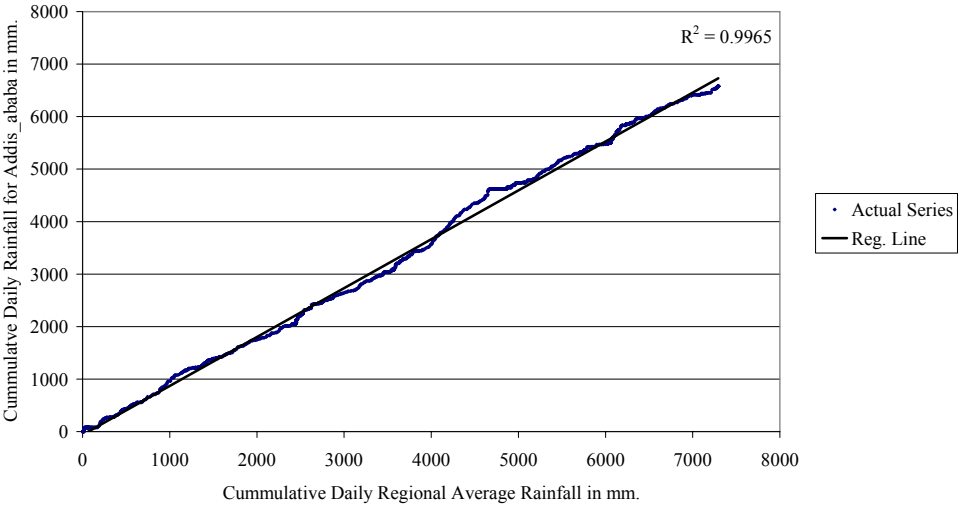

Double Mass Curve for BAHAR_DAR Station
Cummulatve Daily Rainfall for Bahar_Dar in mm.
Cummulative Daily Regional Average Rainfall in mm.
R2 = 0.9975
Actual Series
Reg. Line
0
1000
2000
3000
4000
5000
6000
7000
8000
9000

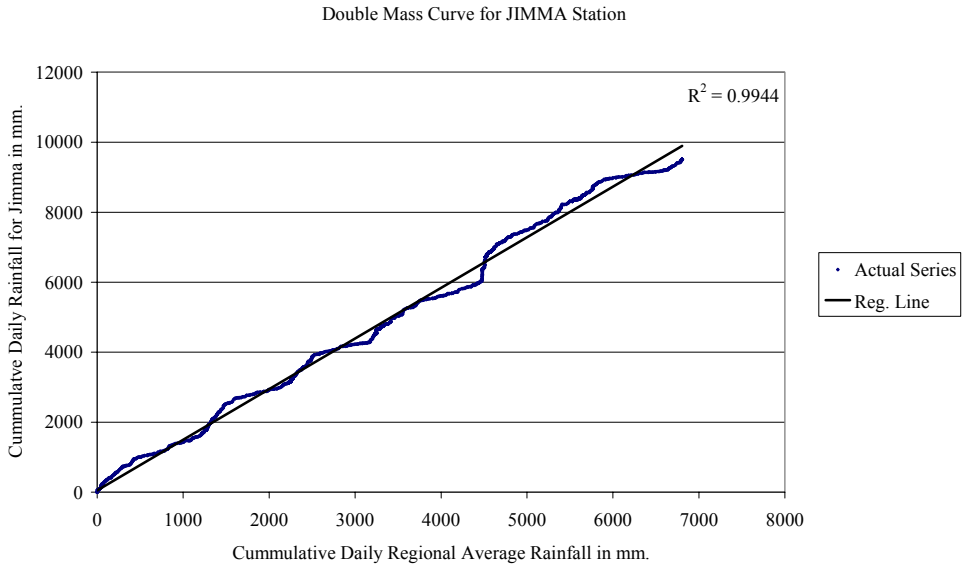

Double Mass Curve for COMBOLCHA Station
Cummulatve Daily Rainfall for DEBRE_MARKOS in mm.
Cummulative Daily Regional Average Rainfall in mm.
R2 = 0.9966
Actual Series
Reg. Line
0
1000
2000
3000
4000
5000
6000
7000
8000

Double Mass Curve for DEBRE_MARKOS Station



Double Mass Curve for ADDIS_ABABA Station




Double Mass Curve for JIMMA Station
$R^2 = 0.9944$
Cummulative Daily Rainfall for Jimma in mm.
12000
10000
8000
6000
4000
2000
0
0
1000
2000
3000
4000
5000
6000
7000
8000
Cummulative Daily Regional Average Rainfall in mm.
Actual Series
Reg. Line

APPENDIX A (4.3)

SOME STATISTICAL PROPERTIES FOR THE RIVER STAGE DATA

Following are some statistical properties for river stage data for:

Wed El-Heleiw station

Year	1993	1994	1995	1996	1997	1998
N	365	340	365	366	365	354
Mean	8.709644	9.065559	9.061014	9.388525	9.405315	9.881017
Median	8.62	8.74	8.71	9.185	9.28	9.43
Std. Error of Mean	0.043681	0.073231	0.052009	0.052457	0.042844	0.081786
Minimum	7.29	7.48	7.93	8.03	8.29	8.32
Maximum	11.12	14.54	12.9	13	11.7	16.5
Std. Deviation	0.834522	1.350321	0.993628	1.003568	0.81854	1.538795
Variance	0.696427	1.823367	0.987296	1.00715	0.670008	2.367891
Skewness	0.584514	1.25222	1.311978	1.005079	0.818647	1.5688
Std. Error of Skewness	0.127689	0.132261	0.127689	0.127516	0.127689	0.129642

Kubur station

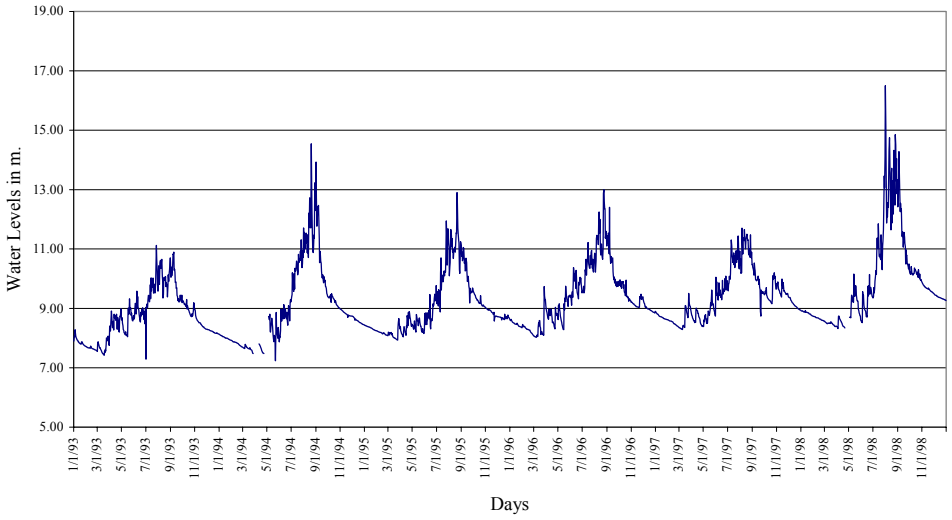
Year	1993	1994	1995	1996	1997	1998
N	291	212	212	271	224	238
Mean	7.589828	8.530283	8.52816	8.352915	8.36433	9.048655
Median	7.02	7.625	7.985	7.88	8.17	8.235
Std. Error of Mean	0.097998	0.147697	0.122454	0.11505	0.09771	0.157067
Minimum	5.71	6.22	6.23	6.23	6.4	0
Maximum	12.38	14.3	12.51	13.62	13.17	14.79
Std. Deviation	1.671718	2.150494	1.78296	1.893964	1.462391	2.423109
Variance	2.794641	4.624623	3.178948	3.587099	2.138588	5.871459
Skewness	0.719946	0.871526	0.49489	0.724035	0.647872	0.386196
Std. Error of Skewness	0.142859	0.167058	0.167058	0.147981	0.162582	0.157788

Kilo_3 station

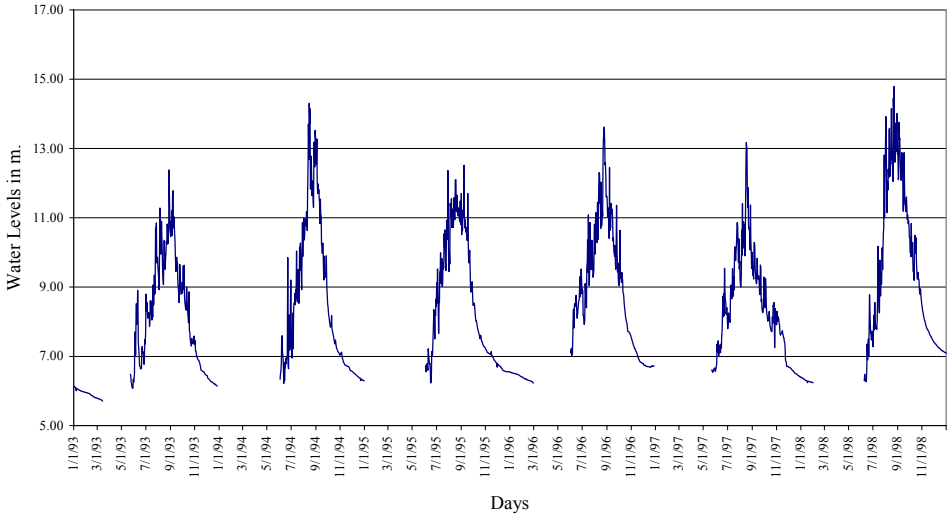
Year	1993	1994	1995	1996	1997	1998
N	142	55	72	127	92	146
Mean	12.25915	12.75564	12.94153	12.85008	12.4588	13.34959
Median	12.15	12.62	13.09	12.68	12.55	13.17
Std. Error of Mean	0.096644	0.183155	0.1098	0.091194	0.097226	0.116319
Minimum	10.12	10.47	11	10.96	10.87	10.64
Maximum	14.53	15.05	14.5	14.84	14.02	15.7
Std. Deviation	1.151645	1.358317	0.931683	1.027708	0.932561	1.405485
Variance	1.326287	1.845025	0.868033	1.056183	0.86967	1.975387
Skewness	0.175792	-0.01704	-0.54831	0.104603	-0.15398	-0.07199
Std. Error of Skewness	0.203429	0.321742	0.282898	0.214848	0.251342	0.200679

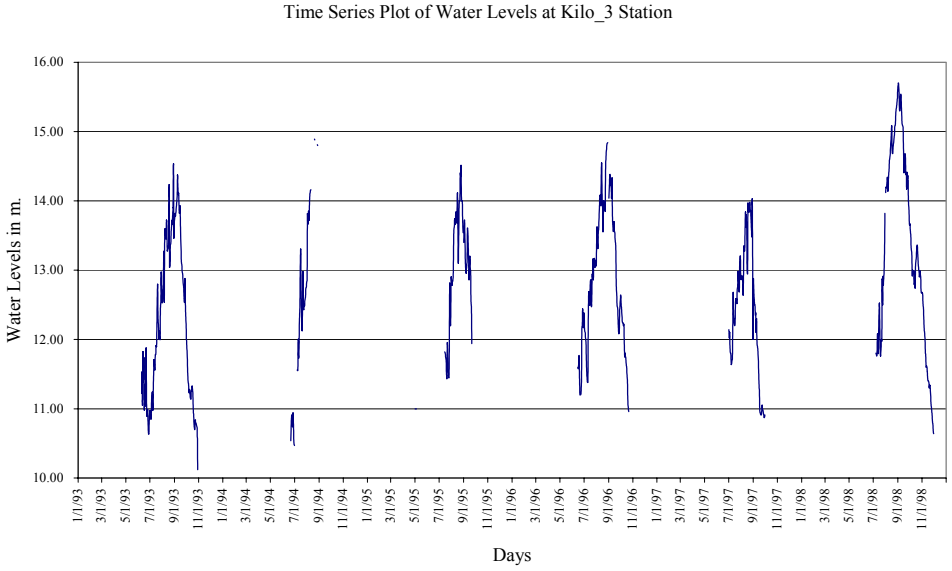
Following are time series plot for river stages for the three stations.

Time Series Plot of Water Level at Wed El-Heleiw Station



Time Series Plot of Water Levels at Kubur Station



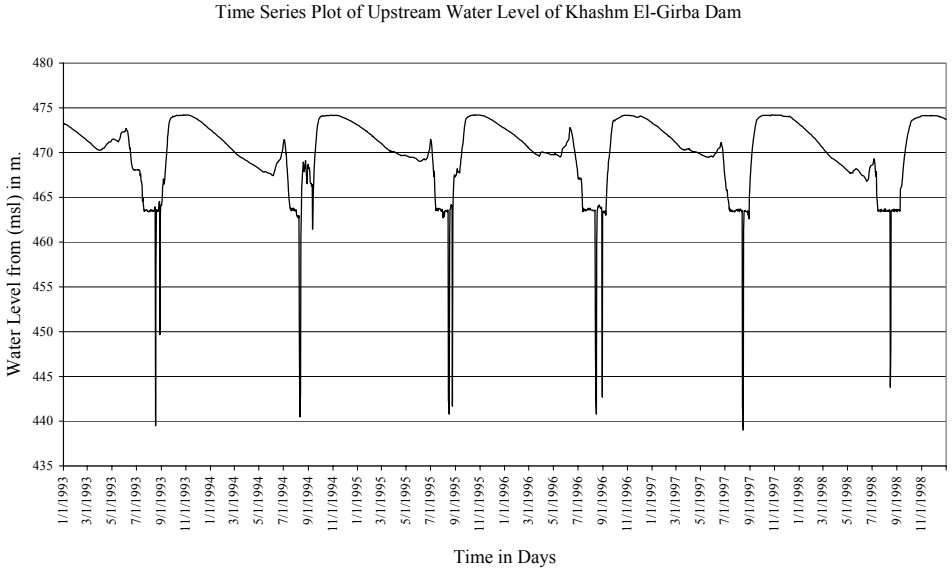

Time Series Plot of Water Levels at Kilo_3 Station
Water Levels in m.
16.00
15.00
14.00
13.00
12.00
11.00
10.00
1/1/93
3/1/93
5/1/93
7/1/93
9/1/93
11/1/93
1/1/94
3/1/94
5/1/94
7/1/94
9/1/94
11/1/94
1/1/95
3/1/95
5/1/95
7/1/95
9/1/95
11/1/95
1/1/96
3/1/96
5/1/96
7/1/96
9/1/96
11/1/96
1/1/97
3/1/97
5/1/97
7/1/97
9/1/97
11/1/97
1/1/98
3/1/98
5/1/98
7/1/98
9/1/98
11/1/98
Days

APPENDIX A (4.4)

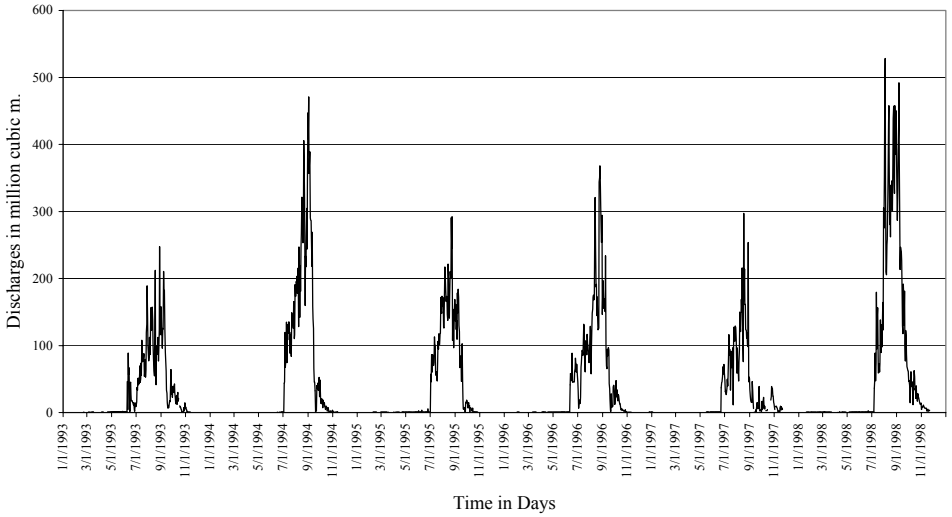
SOME STATISTICAL PROPERTIES FOR Khashm El-Girba RESERVOIR DATA.

	Upstream	Release from the dam	Discharge to the main canal
N	2191	1290	2060
Mean	468.8134	58.05083	4.383
Std. Error of Mean	.2703	2.57420	0.04905
Minimum	373.21	.000	0.0
Maximum	638.50	528.290	53.4
Std. Deviation	12.6510	92.45651	2.226
Variance	160.049	8548.207	4.956
Skewness	-4.979	2.139	5.273
Std. Error of Skewness	0.052	0.068	0.054

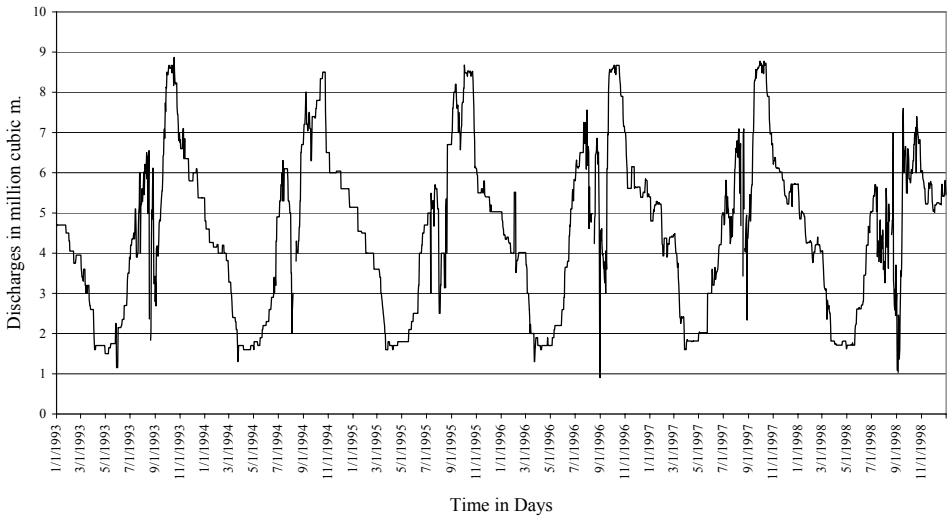
Following are time series plot for the data:



Hydrograph of Khashm El-Girba Dam Release



Hydrograph of Discharges of Khashm El-Girba Dam into the Main Canal



APPENDIX A (4.5)

FILLING IN OF MISSING STAGE AMOUNTS

Name of the station: Kilo_3.

Location: about three kilometers form the of Atbara river with the Nile River.

Period of record: 10 June to 31 October.

Number of years: 6 years.

Percentage of available data points: 72%

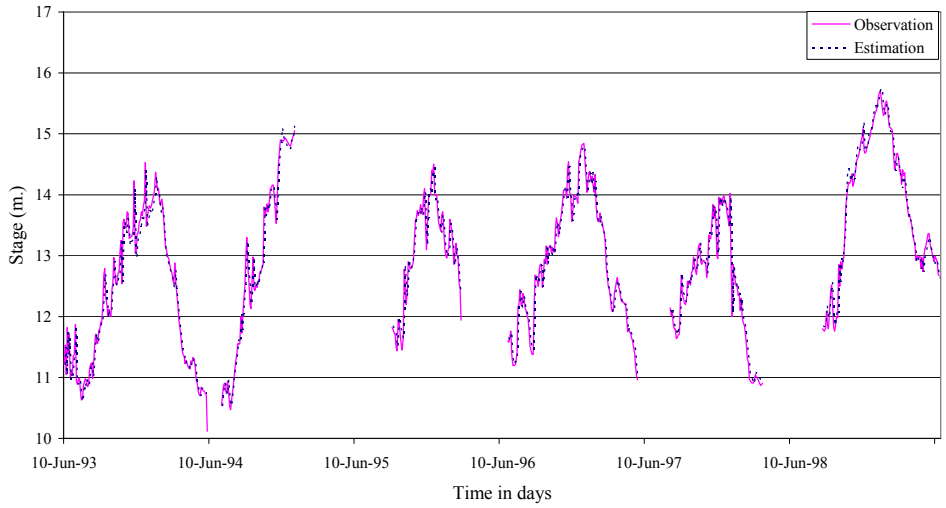
Number of days with missing data: 242.

Three different models were tried using the information available at Kubur and Wed El-Heleiw stations. The results of these models are as follows:

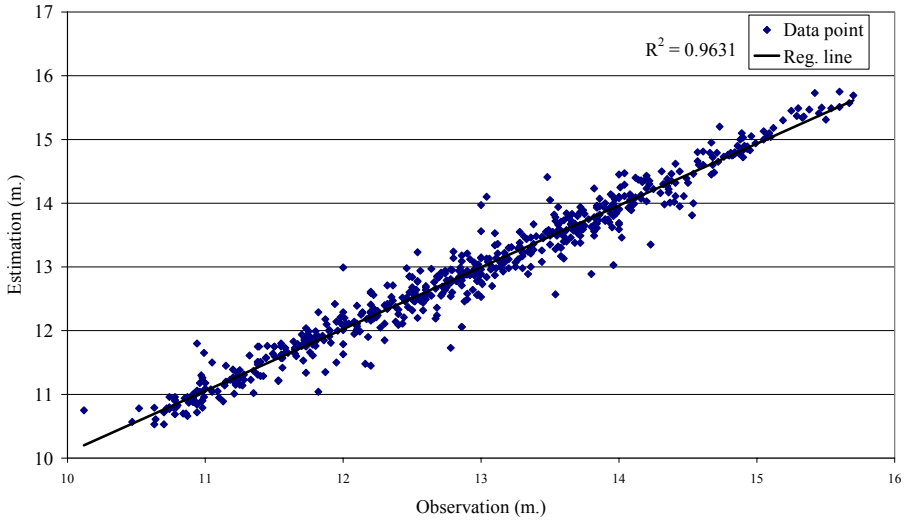
		Model 1	Model 2	Model 3
Regression	Eff (%)	74.19	49.76	96.30
	s.e. of estimate	0.626	0.873	0.235
	D.O.F	2	2	2
	F value	891.311	1007.436	8196.957
Lack of Fit	Sum of Squares	241.748	471.617	34.022
	D.O.F	617	618	611
	Mean Square	0.392	0.763	0.056
	F value	0.436	0.850	0.648
Pure Error	Sum of Squares	0.898	0.898	0.258
	D.O.F	1	1	3
	Mean Square	0.8978	0.8978	0.08595
	constant k	5.545	-	-
	s.e.	0.229	-	-
	t value	24.196	-	-
	coefficient a	0.547	0.416	0.931
	s.e.	0.027	0.037	0.010
	t value	20.216	11.261	97.808
	coefficient b	0.175	0.819	0.084
	s.e.	0.037	0.035	0.012
	t value	4.745	23.110	7.258

Where: Model 1 is the linear relation with constant, Model 2 is the linear relation without constant, and Model 3 is the autoregression relation without constant.

Estimated and observed stages time series plot for Kilo_3 station



Scatterplot of the estimated vs. the observed river stages for Kilo_3 station



APPENDIX A (4.6)

RESULTS OF RATING RELATIONSHIPS DEVELOPMENT

Nonlinear regression analysis was carried to optimize rating relationship's parameters for the three station considered in the study. A detail of these analyses was shown for year 1993 for station Wed El-Heliew, while the rest of the analysis results were summarized showing the parameter's estimates. The analyses were performed using SPSS software.

Analysis Results for Wed El-Heliew Station:-

Following are the results of Nonlinear Regression for data of year 1993

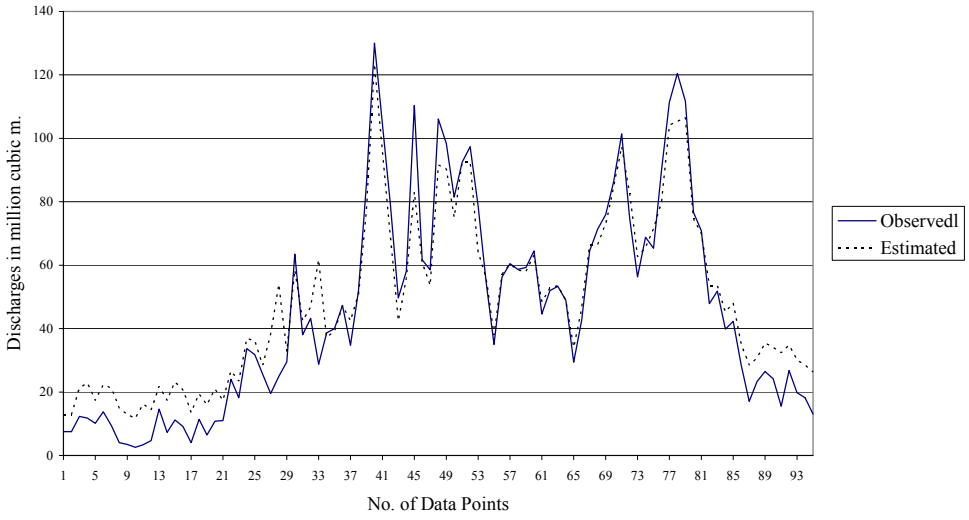
Analysis of Variance

Source	DOF	Sum of Squares	Mean Square
Regression	3	292005.57655	97335.19218
Residual	92	7974.59003	86.68033
Uncorrected Total	95	299980.16659	
Corrected Total	94	99699.53347	

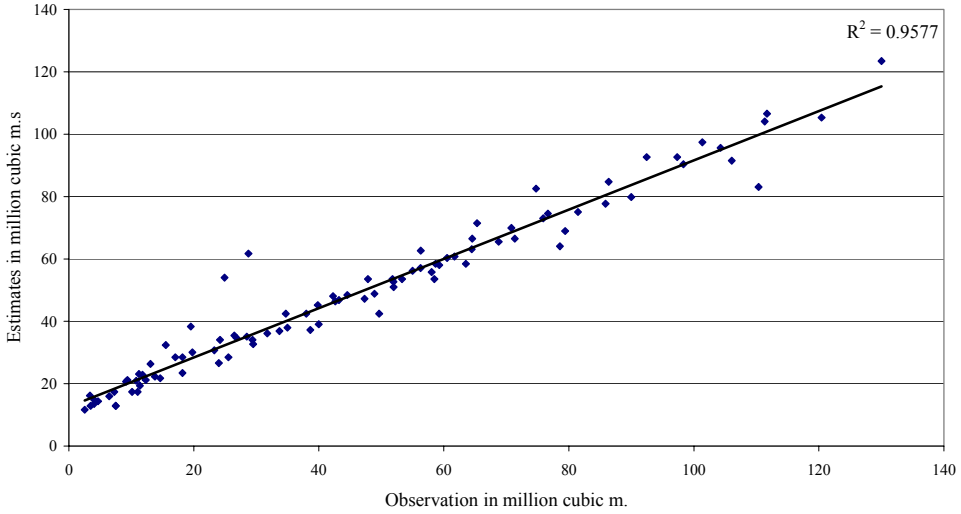
Parameters statistics

Parameter	Estimate	Std. Error	95% Conf. Bounds		95% Range	Trimmed
			Lower	Upper	Lower	Upper
B1	7.3183	.1351	7.0478	7.5887	6.9663	7.5353
B2	7.2900	0.00	7.2900	7.2900	7.2900	7.2900
B3	2.1000	0.00	2.1000	2.1000	2.1000	2.1000

Observed and Estimated Discharges Sequential Plot for Year 1993



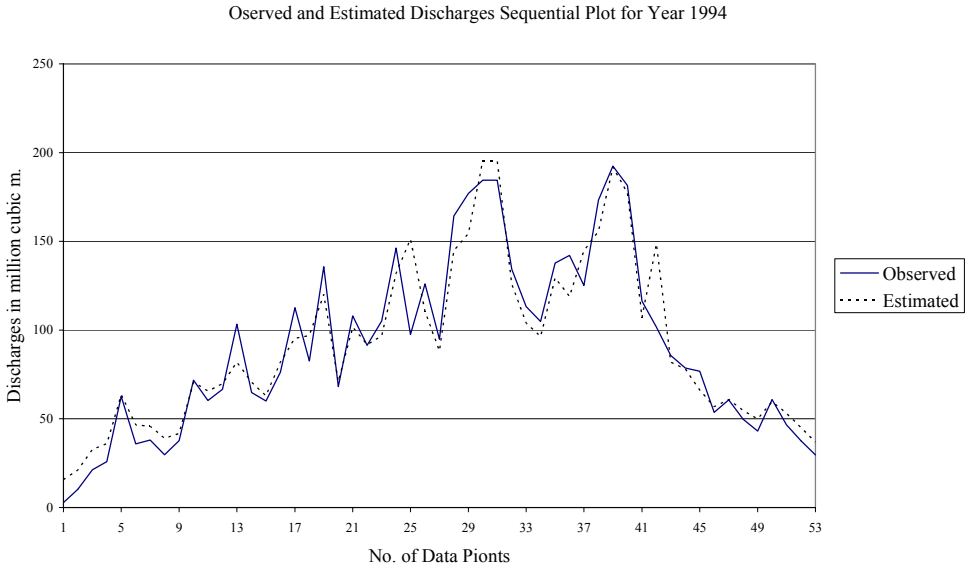
Observed and Estimated Discharges Scatter Plot Year 1993

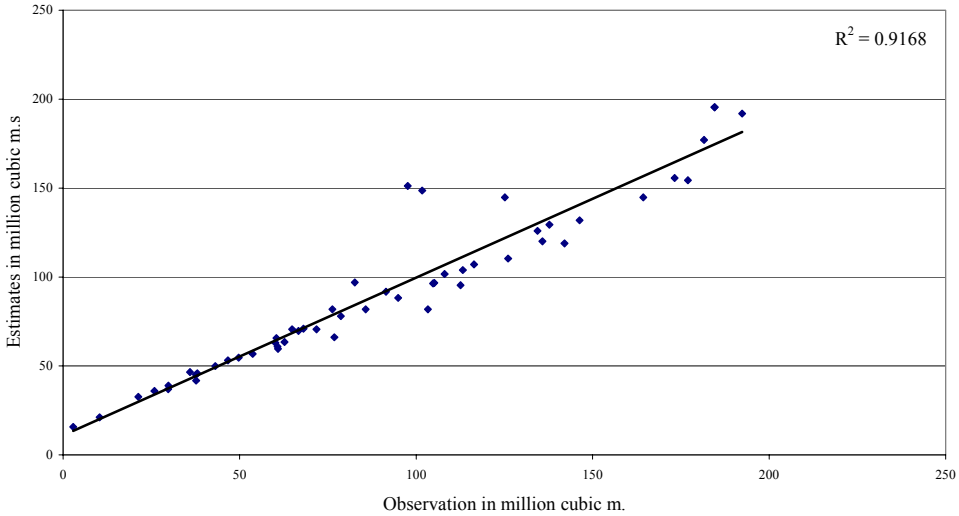


Summary of regression parameters

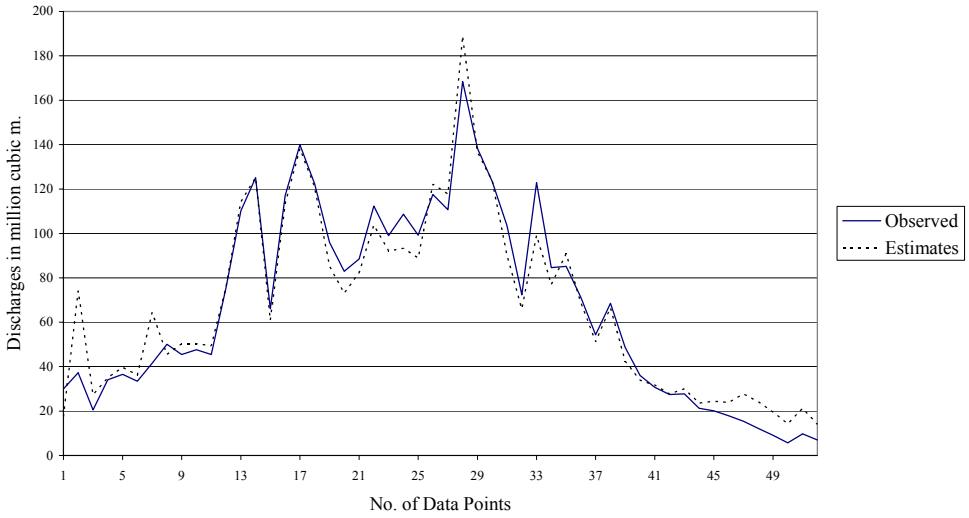
b1(Stage-b2)**b3	Parameter	93	94	95	96	97	98*
Wed ElHeliew	**B1**	7.32	7.54	8.76	8.66	7.91	
	B2	7.29	7.24	7.93	8.03	8.29	
	B3	2.10	1.97	2.10	2.11	2.25	

***= use same parameters of the previous year**

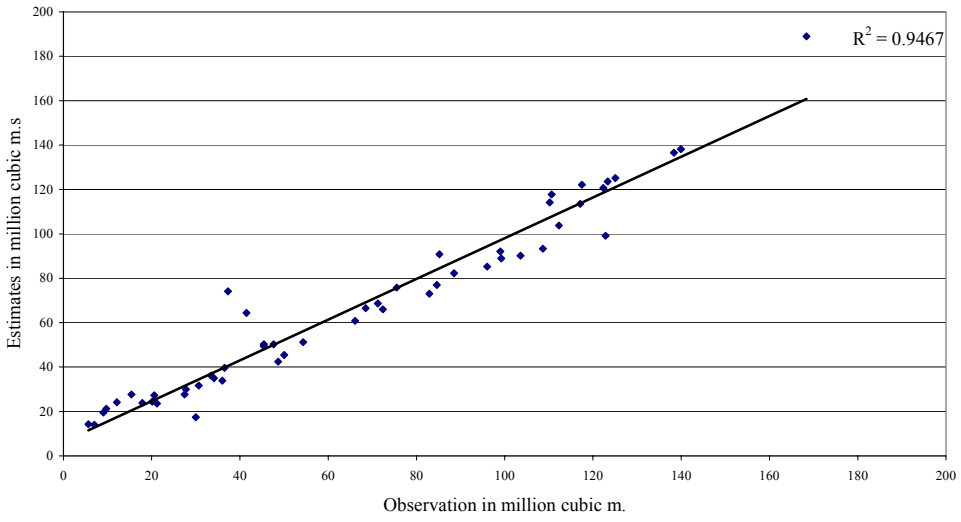




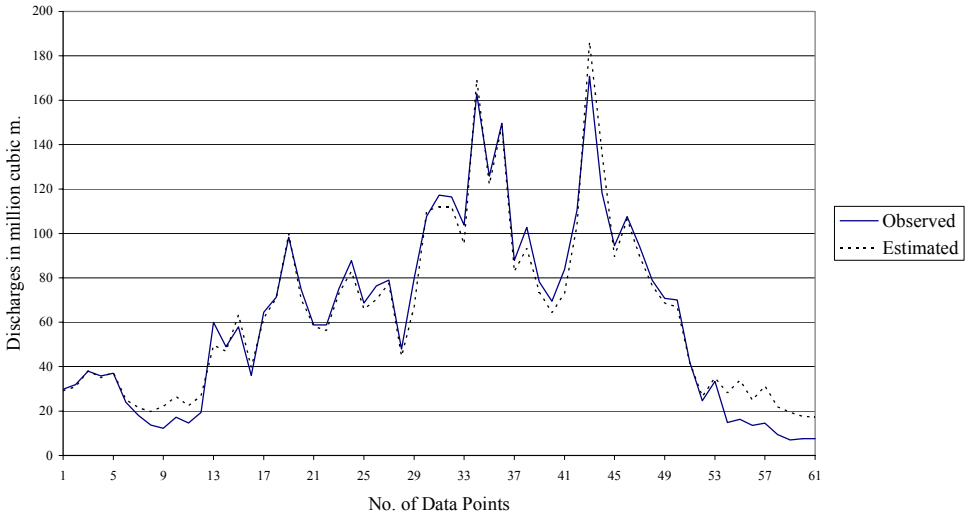
Observed and Estimated Discharges Sequential Plot for Year 1995



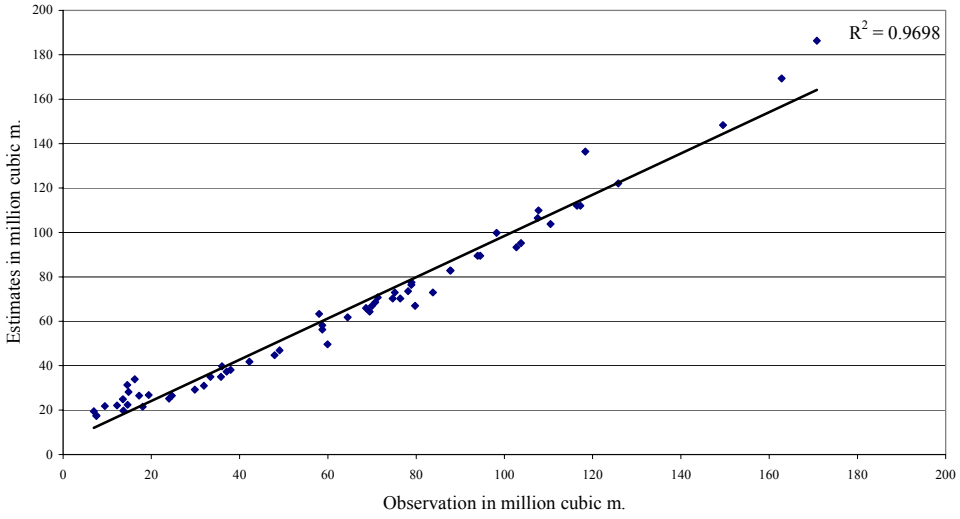
Observed and Estimated Discharges Scatter Plot for Year 95



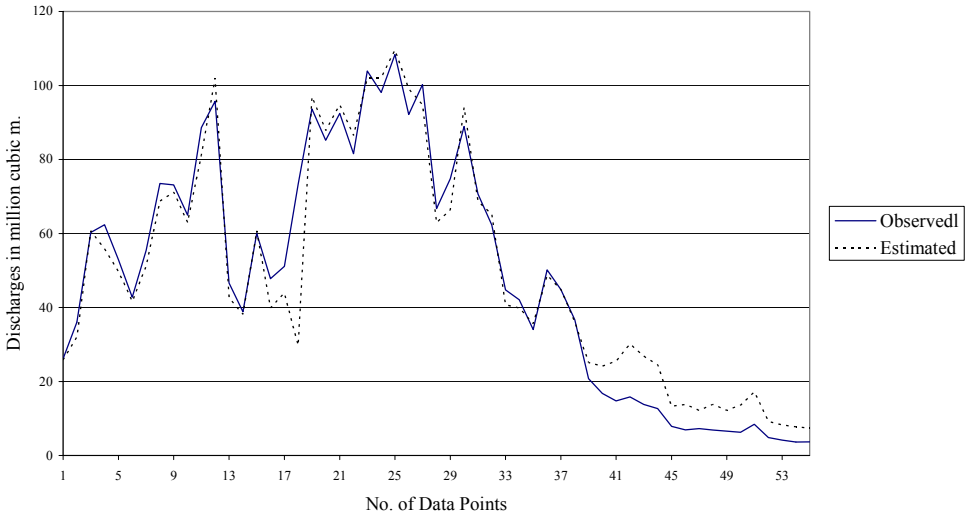
Observed and Estimated Discharges Sequential Plot for Year 1996



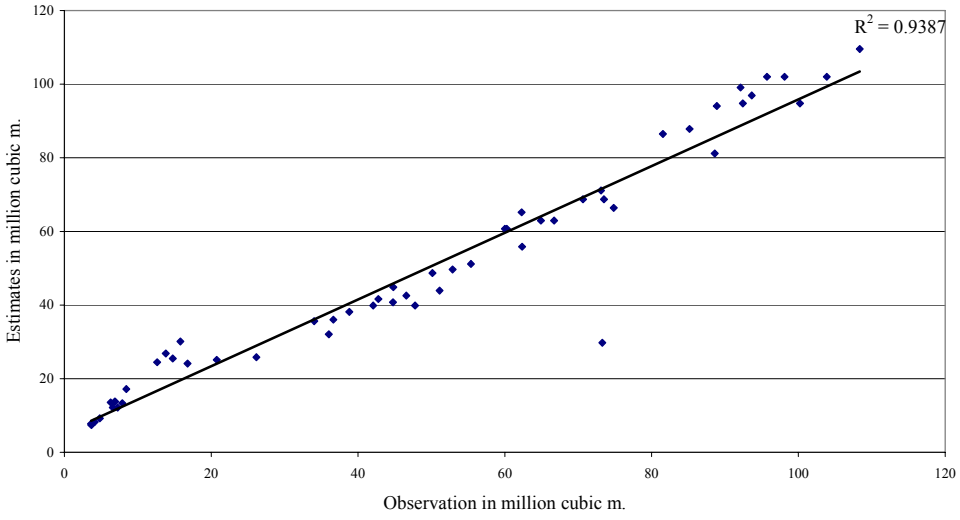
Observed and Estimated Discharges Scatter Plot for Year 96



Observed and Estimated Discharges Sequential Plot for Year 1997



Observed and Estimated Discharges Scatter Plot for Year 97

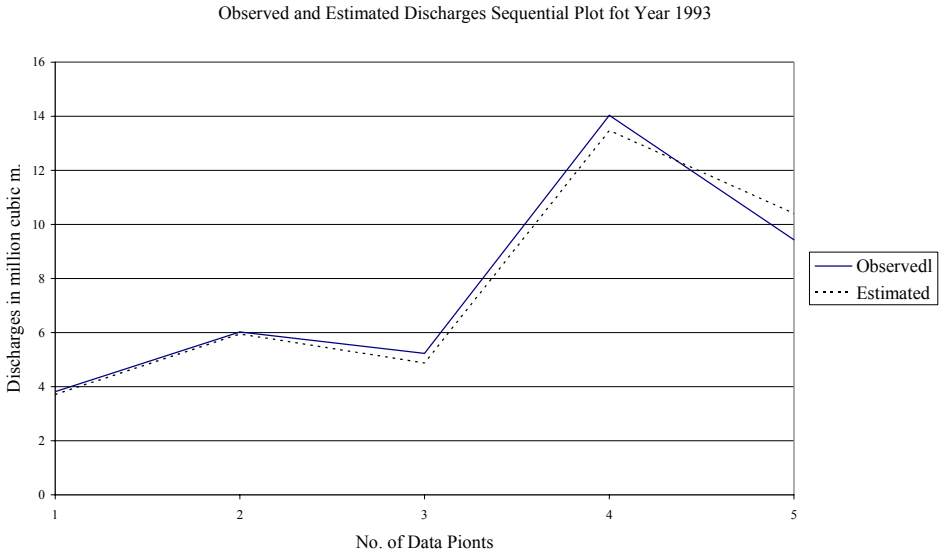


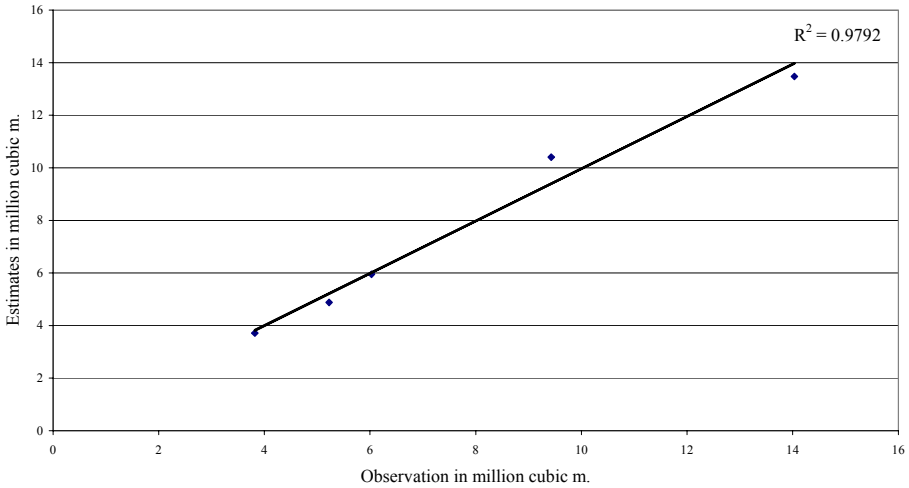
Kubur station:

Summary of regression parameters

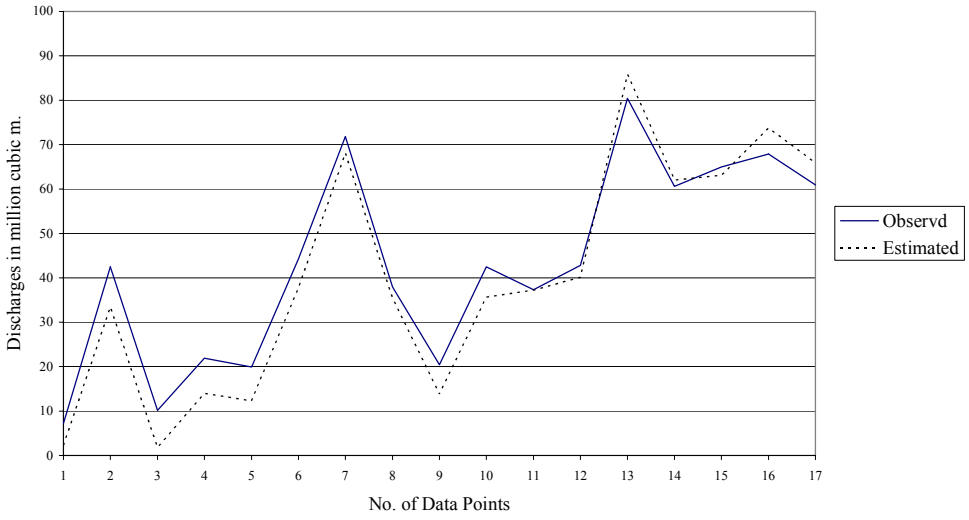
	Parameter	93	94	95	96*	97	98*
Kubur	**B1**	1.96	1.89	1.98		1.74	
	B2	5.71	6.22	6.23		6.24	
	B3	1.82	2.40	2.27		2.28	

***= use same parameters of the previous year**

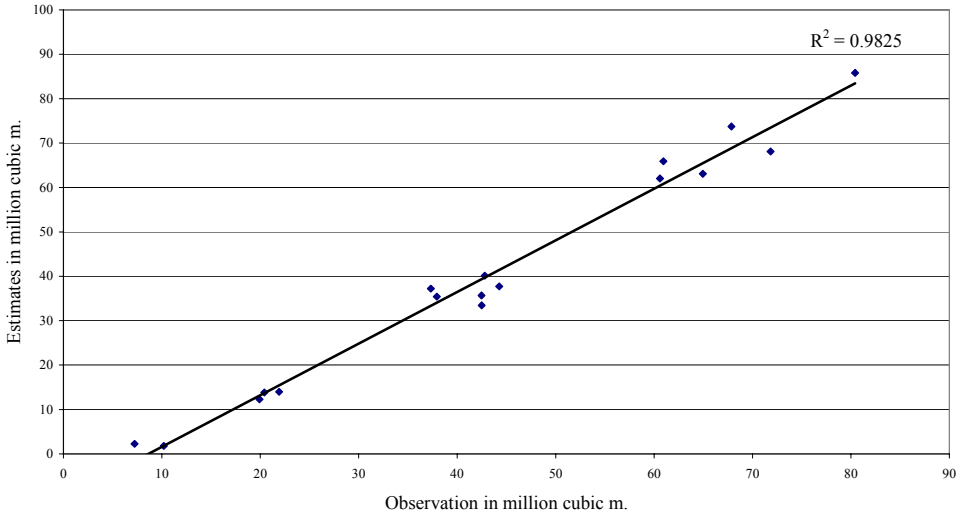




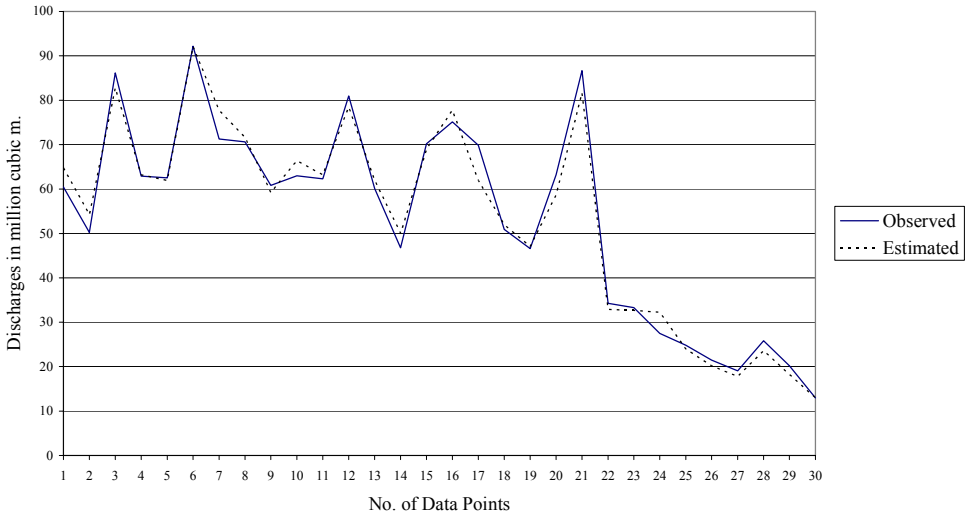
Observed and Estimated Discharges Sequential Plot for Year 1994



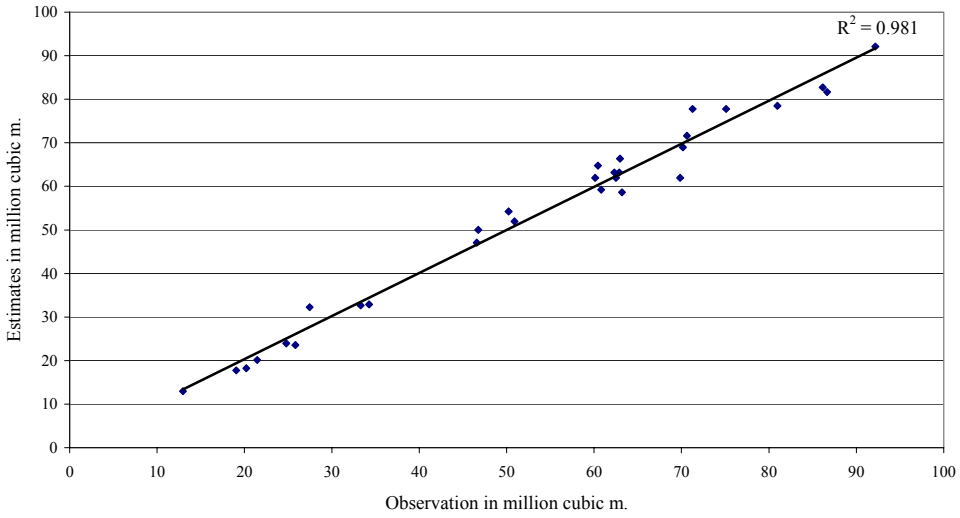
Observed and Estimated Discharges Scatter Plot for Year 1994



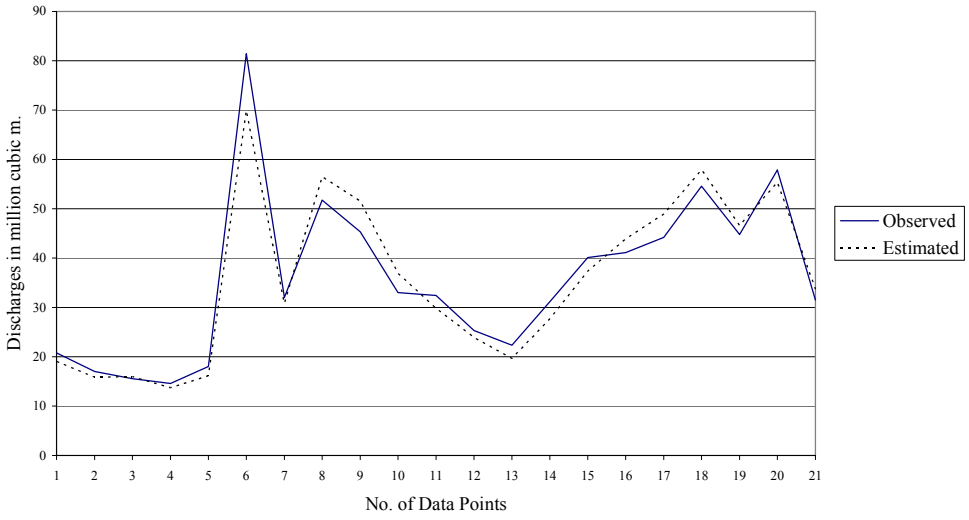
Observed and Estimated Discharges Sequential Plot for Year 1995



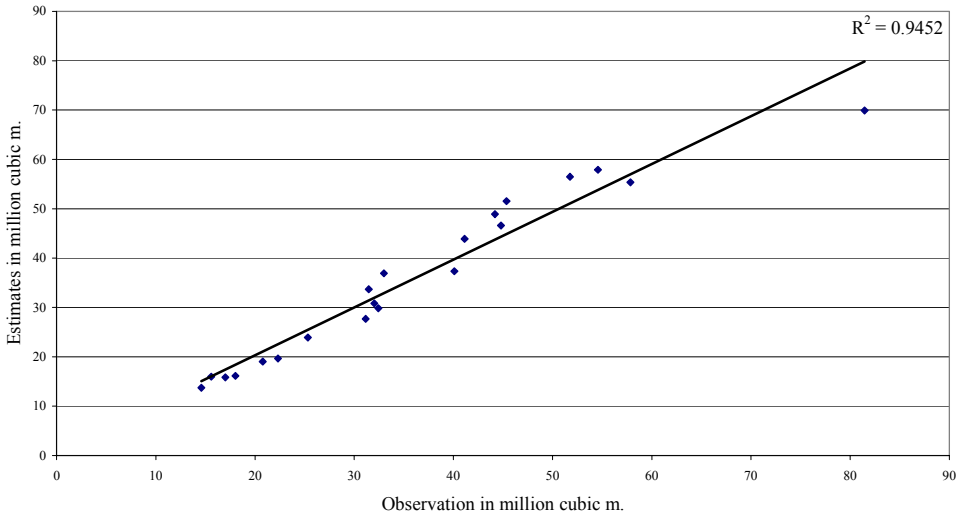
Observed and Estimated Discharges Scatter Plot for Year 1995



Observed and Estimated Discharges Sequential Plot for Year 1997



Observed and Estimated Discharges Scatter Plot for Year 1997



Kilo_3 station:-

Summary of regression parameters

Parameters	1993-1998
B1	0.122
B2	8.269
B3	4.128